QCD matrix elements and truncated showers

Höche, S ; Krauss, F ; Schumann, S ; Siegert, F

Abstract: We derive an improved prescription for the merging of matrix elements with parton showers, extending the CKKW approach. A flavour-dependent phase space separation criterion is proposed. We show that this new method preserves the logarithmic accuracy of the shower, and that the original proposal can be derived from it. One of the main requirements for the method is a truncated shower algorithm. We outline the corresponding Monte Carlo procedures and apply the new prescription to QCD jet production in e+e− collisions and Drell-Yan lepton pair production. Explicit colour information from matrix elements obtained through colour sampling is incorporated in the merging and the influence of different prescriptions to assign colours in the large NC limit is studied. We assess the systematic uncertainties of the new method.

DOI: https://doi.org/10.1088/1126-6708/2009/05/053

ZU-TH 02/09
IPPP/09/14
DCPT/09/28
HD-THEP-09-2
MCNET/09/04

QCD matrix elements and truncated showers

Stefan Höche[1], Frank Krauss[2], Steffen Schumann[3], Frank Siegert[2]

[1] Institut für Theoretische Physik, Universität Zürich, CH-8057 Zürich, Switzerland
[2] Institute for Particle Physics Phenomenology, Durham University, Durham DH1 3LE, UK
[3] Institut für Theoretische Physik, Universität Heidelberg, D-69120, Heidelberg, Germany

Abstract: We derive an improved prescription for the merging of matrix elements with parton showers, extending the CKKW approach. A flavour-dependent phase space separation criterion is proposed. We show that this new method preserves the logarithmic accuracy of the shower, and that the original proposal can be derived from it. One of the main requirements for the method is a truncated shower algorithm. We outline the corresponding Monte Carlo procedures and apply the new prescription to QCD jet production in e^+e^- collisions and Drell-Yan lepton pair production. Explicit colour information from matrix elements obtained through colour sampling is incorporated in the merging and the influence of different prescriptions to assign colours in the large N_C limit is studied. We assess the systematic uncertainties of the new method.

1 Introduction

With the LHC becoming fully operational in the near future, searches for new physics beyond the Standard Model (SM) will enter a new stage. Despite all optimism, a majority of the signals currently discussed suffers from severe backgrounds, among them many related to the production of heavy SM particles, such as the weak gauge bosons or top quarks, accompanied with jets. Therefore, it is a central issue for many experimental analyses to correctly describe the production of these particles in conjunction with additional jets. In many cases, the method of choice is to employ simulation programs. In the past, such event generators have proved to be extremely useful and versatile tools, being well capable to describe comparably simple event topologies at sufficiently high precision. However, especially, when additional hard jets complicate the overall event structure, a proper simulation is far from being trivial. Typically such extra radiation is approximated with leading-logarithmic accuracy through the probabilistic description provided by the parton-shower approach. With rising precision needs, however, improved methods become mandatory, which describe the radiation of additional particles beyond the leading-logarithmic approximation employed by the parton showers.

The most traditional of these methods consists in reweighting QCD emissions as described by the parton shower with respective exact matrix elements expressed through parton-shower variables [1]. Unfortunately, the applicability of this method, however elegant, is constrained to cases, where the parton-shower expression either exceeds the matrix element or can be modified accordingly, without hampering the event generation efficiency too badly. This limits the method to a few cases (such as the production of a gluon in $e^+e^- \to q\bar{q}$, top-quark decay plus emission of an additional gluon, or the production of vector bosons in hadron collisions).

In the past years, new and powerful methods for the systematic inclusion of higher order effects into event generation have been developed. In fact, they can be seen as a major theoretical improvement in the detailed understanding of complicated event topologies. The first of these new methods provides means to consistently match NLO calculations for specific processes with the parton shower and has been incorporated into the MC@NLO program [2]. Its basic idea is to organise the counter-terms necessary to technically cancel real and virtual infrared divergences in such a way that the first emission of the parton shower is recovered. This allows the generation of hard kinematics configurations, which can eventually be fed into a parton-shower Monte Carlo. Several applications of the original approach to different processes have been presented, see for instance [3]. A further improvement, aiming at an enhanced independence of both the specific hard process and the details of the parton-shower algorithm is provided by the POWHEG-method [4], which uses the ratio of the actual real radiation matrix element and the original leading-order one to generate the hardest emission. In its current formulation, this approach can be understood as a hybrid of the traditional parton-shower reweighting and the MC@NLO-method. It has been implemented for various processes, see for example [5].

An alternative approach, aiming at an improved description of multi-jet topologies, has been described in [6, 7], and is often called the CKKW algorithm. It has been studied, e.g., in [8, 9] in the cases of W and Z production and the production of pairs of these bosons at the Tevatron and the LHC. The idea here is to separate the phase space for parton emission into two domains, a hard region of jet production and a softer regime of intra-jet evolution. This separation is achieved through a $k_\perp$-type jet measure [10, 11]. Then matrix elements for different parton multiplicities are used to describe the production of a corresponding number of jets, whereas the parton shower is constrained such that it does not produce any additional jets. Leading higher-order effects are added to the various matrix elements by reweighting them with appropriate Sudakov form factors and with ratios of the strong coupling α_s, taken at the $k_\perp$-scales of the individual jet emissions. Independence of the overall result at logarithmic accuracy on the cut in the jet measure is achieved by the interplay of Sudakov form factors and the vetoed parton shower with suitable starting conditions. The method is one of the cornerstones of the event generator SHERPA [12]. A similar approach, formally equivalent in e^+e^- annihilations into hadrons and often called CKKW-L, has been formulated in [13]. In this case a dipole shower rather than a more traditional parton cascade is used to describe QCD radiation beyond fixed order. Recently, first steps towards a merging at NLO precision of the matrix elements in this approach have been presented in [14]. Another, more simplified method, often called the MLM-prescription, has been introduced in [15]. In [16] the approach was applied for the case of top-quark production at hadron colliders. An extension of the algorithm to account for hard QCD radiation in the production of heavy coloured states as they appear in various beyond the Standard Model scenarios was presented in [17]. The differences between the three merging prescriptions have been investigated using the example of W-production at the Tevatron and the LHC [18]. Although the three methods are different and the formal relationship of the latter method with the two former ones could not be completely worked out there, the results presented in these studies are in astonishing agreement. First steps towards such a more formal comparison between the three algorithms has been attempted in [19]. Recently, so-called "multiplicative matching" methods [20] have also attracted some attention.

In this paper we aim at establishing a general theory framework, in which different merging algorithms can be compared on formal grounds. As a result of this we also describe a merging algorithm that preserves the logarithmic accuracy of the shower. We prove this in the most general case, including initial state QCD particles. Although the prescriptions above deal with this case, up to now no statements of the formal accuracy beyond the most leading logarithms has been made in the literature.

Therefore, the outline of this paper is as follows: Section 2 introduces the procedures necessary to consistently evolve parton showers from matrix elements of arbitrary final-state multiplicity, which is one of the key ingredients of the envisaged merging approach. Section 3 presents the general merging procedure. It also sets the theoretical background of the new approach and introduces the event generation algorithm. It is viable for, in principle, arbitrary parton-shower algorithms. We show that the new procedure exactly reproduces the logarithmic accuracy of the shower. The relationship of the framework presented here with the other three algorithms named above is discussed in Appendix A.1, where we show that the original CKKW approach as well as the CKKW-L algorithm can be derived as special cases. We also present results of our algorithm in two relevant cases, namely e^+e^- annihilations into hadrons at LEP I and Drell-Yan-like production of gauge bosons at the Tevatron. In Sec. 5 we describe the Monte Carlo programs, which are employed in this study. Its details, mainly concerning the systematics of the merging itself are presented in Section 6 and discussed in some detail. We conclude in Sec. 7.

2 Prerequisites for merging matrix elements and parton showers

Merging matrix elements with parton showers combines two essentially different approaches to perturbative QCD. Hard matrix elements are exact at some fixed perturbative order in the strong coupling α_s and are therefore efficient in describing exclusive events with fixed jet multiplicity, taking into account non-trivial interferences between different amplitudes. Parton showers are employed to generate the QCD radiation pattern, especially at lower scales, close to the hadronisation scale Λ_{QCD}. Their application resums potentially large logarithmic corrections due to Bremsstrahlung effects. In their description they naturally cannot take into account all interferences, although steps towards including more quantum mechanical effects are currently being discussed, cf. [21]. In a simulation, best results can be expected, if the two approaches are combined consistently, such that each of them operates in those regions of phase space that it describes best.

This necessitates that parton showers can evolve from parton configurations which are given by high multiplicity matrix elements at certain points in phase space and colour space. In such cases the starting conditions for parton showers are often ambiguous. We will see in Sec. 3 that for running the shower in merged samples, it is necessary to interpret the matrix element as a core process and a series of predefined shower branchings. The reason is that the parton-shower evolution can take place at any point in this branching history, giving rise to the truncated shower prescription described in Sec. 2.4. There may be various such histories arising from the same matrix-element configuration. Hence the most appropriate one needs to be identified. Typically, in the spirit of the probabilistic picture underlying the parton shower, it is simply assumed that this is the most probable one.

In this section we derive the formalism and the algorithms necessary to identify the most probable shower history. We also explain the concept of truncated showers and elaborate on various strategies to define colour assignments for the shower input.

2.1 Master evolution equations

To prepare for the following arguments we firstly introduce prototypical evolution equations for parton showers [22], cf. [23] and [24] [1]

$$\frac{\partial g_a(z,t)}{\partial \log(t/\mu^2)} = \int_z^{\zeta_{\max}} \frac{\mathrm{d}\zeta}{\zeta} \sum_{b=q,g} \mathcal{K}_{ba}(\zeta,t)\, g_b(z/\zeta,t) - g_a(z,t) \int_{\xi_{\min}}^{\xi_{\max}} \mathrm{d}\xi \sum_{b=q,g} \xi\, \mathcal{K}_{ab}(\xi,t)\,. \tag{1}$$

In this context, g may denote either a fragmentation function (FF) or a parton distribution function (PDF), cf. for example [26]. The first term on the right hand side encodes resolvable emissions, while the second (related through unitarity) describes unresolved branchings and virtual contributions. The variable t is the evolution parameter, while ζ is the splitting variable of the scheme. The evolution kernels $\mathcal{K}_{ab}$ are obtained from ratios of appropriate $N+1$- and N-particle matrix elements in the respective limit. Schematically

$$\mathcal{K}_{ab}(z,t) \to \frac{1}{\sigma_a^{(N)}(\Phi_N)} \frac{\mathrm{d}^2\sigma_b^{(N+1)}(z,t;\Phi_N)}{\mathrm{d}\log(t/\mu^2)\,\mathrm{d}z}\,. \tag{2}$$

Here Φ_N denotes the respective N-particle phase space configuration, which does not play a role for the limiting behaviour of $\sigma_b^{(N+1)}(z,t;\Phi_N)$. Equation (2) conversely implies that any splitting kernel $\mathcal{K}_{ab}$ can be substituted by an appropriate ratio of matrix elements, because respective differences are always subleading. For the most common case of standard DGLAP evolution [27], the kernels are easily identified through

$$\mathcal{K}_{ab}(z,t) \to \frac{\alpha_s(z,t)}{2\pi} P_{ab}(z)\,, \tag{3}$$

with $P_{ab}(z)$ being the standard DGLAP splitting functions. If Eq. (1) is written in inclusive form, i.e. $\xi_{\min} \to 0$, $\xi_{\max}, \zeta_{\max} \to 1$, the last term vanishes because of momentum sum rules for the kernels. In exclusive form, where the ζ- and ξ-boundaries are determined by a resolution criterion for parton emission,

[1] The functions g can also explicitly depend on the splitting variable ζ, like in the case of angular ordered DGLAP evolution [25]. This does, however, not complicate our formalism because corresponding terms do not modify the form of the Sudakov form factor. As far as such an evolution is concerned, the corresponding notation in the resolved term of Eq. (1) is implicit.

it can be written as the logarithmic derivative of the Sudakov form factor [28]

$$\Delta_a(\mu^2, t) \;=\; \exp\left\{ -\int_{\mu^2}^{t} \frac{\mathrm{d}\bar{t}}{\bar{t}} \int_{\xi_{\min}}^{\xi_{\max}} \mathrm{d}\xi \sum_{b=q,g} \frac{1}{2}\, \mathcal{K}_{ab}(\xi, \bar{t}) \right\} . \tag{4}$$

The factor $1/2$ is equivalent to ξ under the integral and the sum over parton species and avoids double-counting identical decay channels. Equation (4) has the generic form of a no-emission probability between the scales μ^2 and t. Potential differences in various shower algorithms implementing QCD evolution arise due to different evolution kernels $\mathcal{K}$ or a different interpretation of the evolution and splitting variables. This in turn corresponds to the choice of a factorisation scheme. If Eq. (1) is modified accordingly, the kernels can also incorporate more than two partons, for example in the case of dipole [29, 30, 31, 32] or dipole-like [33, 34, 35] cascades. Therefore, in this paper the term "parton shower" also includes the dipole or dipole-like shower algorithms.

2.2 Branching probabilities

Equation (1) is commonly rewritten in terms of g/Δ. The corresponding form reads

$$\frac{\partial}{\partial \log(t/\mu^2)} \frac{g_a(z,t)}{\Delta_a(\mu^2,t)} = \frac{1}{\Delta_a(\mu^2,t)} \int_z^{\zeta_{\max}} \frac{\mathrm{d}\zeta}{\zeta} \sum_{b=q,g} \mathcal{K}_{ba}(\zeta,t)\, g_b(z/\zeta,t) . \tag{5}$$

This immediately yields the no-branching probabilities for unconstrained (forward) and constrained (backward) shower evolution as [22]

$$\mathcal{P}^{(F)}_{\mathrm{no},\,a}(t,t') = \frac{\Delta_a(\mu^2,t')}{\Delta_a(\mu^2,t)} = \exp\left\{ -\int_t^{t'} \frac{\mathrm{d}\bar{t}}{\bar{t}} \int_{\zeta_{\min}}^{\zeta_{\max}} \mathrm{d}\zeta \sum_{b=q,g} \frac{1}{2}\, \mathcal{K}_{ab}(\zeta,\bar{t}) \right\} , \tag{6}$$

and

$$\mathcal{P}^{(B)}_{\mathrm{no},\,a}(z,t,t') = \frac{\Delta_a(\mu^2,t')\, g_a(z,t)}{\Delta_a(\mu^2,t)\, g_a(z,t')} = \exp\left\{ -\int_t^{t'} \frac{\mathrm{d}\bar{t}}{\bar{t}} \int_z^{\zeta_{\max}} \frac{\mathrm{d}\zeta}{\zeta} \sum_{b=q,g} \mathcal{K}_{ba}(\zeta,\bar{t})\, \frac{g_b(z/\zeta,\bar{t})}{g_a(z,\bar{t})} \right\} . \tag{7}$$

The standard procedure for constructing a parton-shower algorithm is to write the differential branching probability $\mathcal{P}_{\mathrm{branch,\,a}}$ as

$$\mathcal{P}_{\mathrm{branch,\,a}}(t,t') = \frac{\partial \mathcal{P}_{\mathrm{no,\,a}}(t,t')}{\partial \log(t/\mu^2)} . \tag{8}$$

Given the current evolution scale t', at which an actual emission took place, a new scale t for the next parton emission is chosen according to this probability.

2.3 Shower histories from matrix elements

An obvious point in the recursive algorithm sketched above is to define the first or hardest scale, at which the parton shower starts off. In order to obtain such suitable starting conditions for parton showers from arbitrary matrix elements, a clustering algorithm needs to be defined, which corresponds to "running the shower evolution backwards" on the respective matrix element. It identifies how, in a parton-shower picture, the matrix element would have originated from a lower multiplicity matrix element and a shower branching. Applied iteratively, it leads to the definition of a core process, which cannot be further decomposed and a sequence of shower branchings yielding the actual final state. The tasks for the algorithm are thus twofold: Firstly, within an arbitrary n-parton final state the most probable splitting in terms of shower evolution starting with $n-1$ partons needs to be found. Secondly, corresponding partons must be recombined to obtain the respective final state. In order to construct this algorithm, the shower evolution simply has to be "inverted", which gives the following recipe:

The criterion of the cluster algorithm is defined by the shower evolution kernels.
The recombination scheme is given by the inverted shower kinematics.

A side effect of this prescription is that during backward clustering the hard matrix elements, potentially no strict hierarchy is found in the shower evolution parameter. However, such situations can only arise for kinematic configurations of the matrix element, which are beyond the accuracy of the shower. Therefore they do not pose a problem here.

2.4 Truncated showering

Assume a parton-shower history constructed from a matrix element along the lines of Sec. 2.3 and consisting of a core process plus a single additional branching at scale t, which we call matrix-element branching. As a consequence of Eqs. (6) and (7), parton shower emissions may take place at scales $t' > t$. This leads to a situation, where, due to additional partons originating from these branchings, the kinematics of the matrix-element branching at t needs to be redefined. This corresponds to a truncated shower, identical to the situation discussed in [4], where a mismatch between hardest emission, in terms of transverse momentum, and the parton-shower evolution defined in opening angles has been noted and resolved. There the solution to this situation has been coined "truncated showering", because the evolution stops at the lower, dynamical scale t, unlike ordinary shower evolution, which stops at the universal cutoff scale. A prescription to unambiguously reconstruct the kinematics of matrix-element branchings is needed. The most natural choice is to compute the evolution, splitting and angular variables of matrix element branchings in the shower scheme and reconstruct the branching using the shower kinematics, after the final state of the branching at $t' > t$ is fully determined. In other words:

Splitting kernels introduced by Eq. (2) *define evolution, splitting and angular variables in the shower scheme, rather than the kinematics of the corresponding branching.*

This coincides with the fact that such branchings are interpreted as predetermined nodes during shower evolution. It must simply be guaranteed that the evolution stops at the corresponding scale, inserts the node and restarts afterwards. If for any reason (e.g. energy-momentum conservation) the matrix element branching cannot be reconstructed after a truncated shower branching, this shower branching must be vetoed.

2.5 Colour treatment

The treatment of colour is a central issue when dealing with matrix element and shower generation in QCD processes. Shower formulations inherently are correct only at leading order in $1/N_C$, although in [21] first attempts to overcome this limitation have been presented. Therefore, once matrix elements are to be combined with showers, the treatment of colour must be adjusted. A simple and obvious way to do so is to interpret the hard matrix element in the large-N_C limit to define colour partners of splitting partons in the shower language. This problem is more easily solved, when colours are not summed over but if an algorithm is employed, which unambiguously assigns a certain set of colours to the external particles in the hard matrix element. The basic idea is then to sample over colours in a Monte Carlo fashion rather than summing over them.

It was pointed out in [36, 37] that the colour-flow decomposition of QCD amplitudes is advantageous over both the fundamental and (if existent) the adjoint representation decomposition. The key point is that in the colour-flow decomposition each colour octet is treated as a 3×3 index field whose additional degree of freedom is removed by a projector onto the colour-octet subspace. Since this issue is central to all our simulations incorporating fixed colour assignments in the hard matrix element, we briefly recall the basics of the algorithm. As an example we consider an n-gluon amplitude $\mathcal{A}(1,\dots,n)$. This amplitude can be decomposed in the colour-flow basis as [36]

$$\mathcal{A}(1,\dots,n) = \sum_{\sigma\in S_{n-1}} \delta^{i_1\bar{\jmath}_{\sigma_2}}\delta^{i_{\sigma_2}\bar{\jmath}_{\sigma_3}}\dots\delta^{i_{\sigma_n}\bar{\jmath}_1}\, A(1,\sigma_2,\dots,\sigma_n)\ . \tag{9}$$

Here i_{σ_k} and $\bar{\jmath}_{\sigma_k}$ denote the 3- and $\bar{3}$-index of parton σ_k, respectively and the sum runs over all possible permutations of the set $\{2,\dots,n\}$. The quantities $A(1,\sigma_2,\dots,\sigma_n)$ are called colour-ordered or partial amplitudes. They depend on the kinematics of the process only. All information about colour is incorporated in

respective prefactors. Therefore any colour-ordered amplitude only contains planar diagrams, which greatly alleviates the computation. A convenient way to interpret Eq. (9) is to consider it as the decomposition of the full QCD amplitude into subamplitudes in the large-N_C limit. Feeding the results from the matrix-element calculation into a shower program, the corresponding colour connections are thus readily determined if one of the terms in the sum is picked as the most probable colour structure and identify the colour flow according to its colour factor. In this context we use the fact that interference terms between two different colour structures are always subleading in $1/N_C$.[2]

An algorithm to identify the most probable colour structure could thus look as follows (cf. [39])

1. Compute the full matrix element with randomly assigned colours for external QCD partons.
2. Identify all possible permutations $\{1, \vec{\sigma}\}$, which yield a non-zero value of

   $$\delta^{i_1\bar{j}_{\sigma_2}}\delta^{i_{\sigma_2}\bar{j}_{\sigma_3}}\dots\delta^{i_{\sigma_n}\bar{j}_1} \ . \tag{10}$$

   Label them by $\vec{\sigma}_i$ and compute the corresponding partial amplitudes $A(1, \vec{\sigma}_i)$.
3. If $N_{\vec{\sigma}}$ is the number of identified permutations, choose a partial amplitude with probability

   $$P_{\vec{\sigma}_i} = \frac{|A(1,\vec{\sigma}_i)|^2}{\sum_{j=1}^{N_{\vec{\sigma}}}|A(1,\vec{\sigma}_j)|^2} \tag{11}$$

Because of the way potential partial amplitudes are identified in the colour-flow decomposition, this prescription is similar to the following simplified strategy

1. Compute the full matrix element with randomly assigned colours for external QCD partons.
2. Assign colours in the large-N_C limit at random, respecting the actual point in colour space. This translates into two partons only being colour adjacent at large N_C, if they were colour adjacent at finite N_C.
3. Identify the corresponding permutation $\vec{\sigma}$ and compute the partial amplitude $A(1, \vec{\sigma})$. Accept the configuration with probability

   $$P_{\vec{\sigma}} = \frac{|A(1,\vec{\sigma})|^2}{|\mathcal{A}(1,\dots,n)|^2} \tag{12}$$

Naively, the drawback of the latter algorithm is, that potentially many points have to be drawn for the colour assignment at large N_C. In practice, however, this is sufficiently fast compared to evaluating all possible partial amplitudes for a single nontrivial point in colour space. Also, in principle the full amplitude squared, $|\mathcal{A}(1,\dots,n)|$, might be much smaller than the sum of partial amplitudes squared, such that acceptance probabilities are modified. This algorithm is still sufficiently accurate, since respective differences are always subleading in $1/N_C$.

The above algorithm works for arbitrary QCD and QCD-associated matrix elements, since quark pairs can always be reinterpreted as colour octet objects. Matrix-element configurations might exist, which do not allow an immediate projection onto large N_C because of the U(1) pseudo-gluon. In this case, a new point in colour space can safely be assigned, because the respective contribution to the total cross section is subleading.

3 The merging algorithm

The central idea for algorithms merging matrix elements with parton showers is to replace products of splitting kernels related to hard emissions in the shower evolution with the appropriate matrix elements, thus reinstalling information about the full hard process under consideration. Directly implementing a ratio of hard matrix elements in form of a splitting kernel has the apparent disadvantage that the respective

[2] This argument holds in the colour-flow decomposition and the fundamental representation decomposition. For the latter, see for example [38].

phase-space integration proceeds in terms of shower kinematics and is thus hard to optimise in a generic way. A better technique is to first compute the matrix element and then reweight it such that, to the accuracy of the parton shower, the corresponding shower expression is obtained. To pursue this strategy, the corresponding no-emission probabilities of the parton shower, i.e. its Sudakov form factors, must be known. This can, however easily be achieved because they emerge directly from the evolution equations on which the shower is based.

Only one additional ingredient is eventually needed, namely a criterion, which defines how to separate matrix-element and parton-shower domain. It will be shown in Sec. 4 that a general form of this criterion can be found, which is based on the soft and collinear behaviour of QCD at the next-to-leading order. We will refer to it as the "jet criterion". At the present stage, the precise form of the jet criterion is unimportant and it is sufficient and helpful to think of it in an abstract way.

3.1 Construction of the algorithm

The basic idea of the merging – to separate the matrix-element and parton-shower domains through a cut in the emission phase space, defined by a jet criterion – corresponds to a simple phase-space slicing. We therefore define the evolution kernels for the matrix-element and parton-shower domains

$$\mathcal{K}_{ab}^{\mathrm{ME}}(\xi,\bar{t}) = \mathcal{K}_{ab}(\xi,\bar{t})\,\Theta\left[Q_{ab}(\xi,\bar{t}) - Q_{\mathrm{cut}}\right] \qquad \text{and} \qquad \mathcal{K}_{ab}^{\mathrm{PS}}(\xi,\bar{t}) = \mathcal{K}_{ab}(\xi,\bar{t})\,\Theta\left[Q_{\mathrm{cut}} - Q_{ab}(\xi,\bar{t})\right]\,, \tag{13}$$

where Q_{ab} denotes the jet criterion and Q_{cut} is its cut value. Correspondingly, the two partial Sudakov form factors are given by

$$\Delta_a^{\mathrm{ME/PS}}(\mu^2,t) = \exp\left\{-\int_{\mu^2}^{t}\frac{\mathrm{d}\bar{t}}{\bar{t}}\int_{\xi_{\min}}^{\xi_{\max}}\mathrm{d}\xi\sum_{b=q,g}\frac{1}{2}\,\mathcal{K}_{ab}^{\mathrm{ME/PS}}(\xi,\bar{t})\right\}\,. \tag{14}$$

They are related to the full Sudakov form factor, Eq. (4), through

$$\Delta_a(\mu^2,t) = \Delta_a^{\mathrm{PS}}(\mu^2,t)\,\Delta_a^{\mathrm{ME}}(\mu^2,t)\,. \tag{15}$$

In fact, Eq. (15) effectively encodes the complete merging approach. There, ultimately $\mathcal{K}^{\mathrm{ME}}$ will be replaced with a ratio of matrix elements, according to Eq. (2). During the following rewrite of the evolution equations we simply identify, how the factorisation property of Sudakov form factors must then be interpreted and employed for event generation. No further proof for the correctness of the algorithm at any logarithmic order is necessary, because this follows directly from the accuracy implemented in the parton-shower evolution. In other words, the proposed merging scheme does not impair the logarithmic accuracy of the parton shower.

This can be summarised as:

> *When correcting parton-shower evolution kernels through higher-order matrix elements, the master evolution equation, Eq.* (1)*, must be respected. This ensures that the accuracy generated by the factorisation scheme and the parton shower is fully restored.*

We start by defining the conditional backward no-branching probability in the parton-shower domain[3],

$$\tilde{\mathcal{P}}_{\mathrm{no},\,a}^{(B)\,\mathrm{PS}}(z,t,t') = \frac{\Delta_a^{\mathrm{PS}}(\mu^2,t')\,\tilde{g}_a(z,t)}{\Delta_a^{\mathrm{PS}}(\mu^2,t)\,\tilde{g}_a(z,t')} = \exp\left\{-\int_t^{t'}\frac{\mathrm{d}\bar{t}}{\bar{t}}\int_z^{\zeta_{\max}}\frac{\mathrm{d}\zeta}{\zeta}\sum_{b=q,g}\mathcal{K}_{ba}^{\mathrm{PS}}(\zeta,\bar{t})\,\frac{\tilde{g}_b(z/\zeta,\bar{t})}{\tilde{g}_a(z,\bar{t})}\right\}\,. \tag{16}$$

It corresponds to the modified evolution equation

$$\frac{\partial}{\partial\log(t/\mu^2)}\log\frac{\tilde{g}_a(z,t)}{\Delta_a^{\mathrm{PS}}(\mu^2,t)} = \int_z^{\zeta_{\max}}\frac{\mathrm{d}\zeta}{\zeta}\sum_{b=q,g}\mathcal{K}_{ba}^{\mathrm{PS}}(\zeta,t)\,\frac{\tilde{g}_b(z/\zeta,t)}{\tilde{g}_a(z,t)}\,. \tag{17}$$

Equation (16) differs with respect to the standard parton-shower evolution because of the Θ-function, restricting emissions to $Q < Q_{\mathrm{cut}}$, i.e. the soft and collinear domain. Its interpretation is therefore straightforward and gives a rule for the modified shower algorithm in the merging:

[3]From here on we focus on backward evolution. The corresponding reasoning for forward evolution follows trivially.

Standard parton-shower evolution is implemented, but radiation with $Q > Q_{\rm cut}$ is vetoed.

Note that for the case of an initial-state parton shower, typically described with a backward evolution, the initial scale of the PDF's is set by the core process of the event.

If Eq. (16) is employed as is, including the phase-space restriction, the newly defined functions $\tilde{g}$ do not obey the same evolution as the original functions g. Factorisation is thus violated. If we want the two evolutions to agree, we have to guarantee that the full no-branching probability in the merging approach is given by Eq. (7). This leads to the definition of the no-emission probability in the matrix-element domain according to

$$\mathcal{P}^{(B)\,\rm ME}_{{\rm no},\,a}(t,t') = \frac{\mathcal{P}^{(B)}_{{\rm no},\,a}(z,t,t')}{\mathcal{P}^{(B)\,\rm PS}_{{\rm no},\,a}(z,t,t')} = \frac{\Delta^{\rm ME}(\mu^2,t')}{\Delta^{\rm ME}(\mu^2,t)}\,, \quad \text{where} \quad \mathcal{P}^{(B)\,\rm PS}_{{\rm no},\,a}(z,t,t') = \frac{\Delta^{\rm PS}_a(\mu^2,t')\,g_a(z,t)}{\Delta^{\rm PS}_a(\mu^2,t)\,g_a(z,t')}\,. \tag{18}$$

It is interesting to note that $\mathcal{P}^{(B)\,\rm ME}_{{\rm no},\,a}$ is independent of z, which effectively is an outcome of the factorisation properties of PDF's and FF's.

If we assume that a "most probable" shower history of the hard matrix element is identified through a backward-clustering algorithm, which employs the clustering criterion and the recombination scheme of the shower, cf. Sec. 2.3, then we obtain the rule:

The weight, Eq. (18)*, is assigned to any leg with production scale t' and decay scale t found during backward clustering. Strong couplings are evaluated at the nodal scales of parton recombination.*

The reasoning is easily explained. Hard matrix elements in the factorisation scheme of the shower have the same limiting behaviour as the splitting kernels $\mathcal{K}$, once colour-adjacent partons become close in phase space. Backward clustering will identify a hierarchical structure for the factorisation of hard matrix elements into lower-multiplicity matrix elements and splitting kernels. Eventually, a core process is found, which cannot be further decomposed and which corresponds to the starting conditions for a respective shower evolution. Matrix elements, however, do not implement the no-branching probabilities generated by parton showers. Also the strong coupling is evaluated at a common scale, rather than the nodal scales of splittings. Corresponding corrections must therefore be implemented.

An immediate consequence of the above defined algorithm is that the total cross section of merged samples can only be influenced by the difference between full hard matrix elements and the corresponding product of splitting kernels times the core process. In this respect, we obtain the rule:

To compute hadronic cross sections, PDF's must be evaluated at the scale of the core process defined through backward clustering.

This prescription is independent of the multiplicity of the matrix element, because showering always starts at the scale of the core process, cf. Eq. (7). A mismatch in the two scales would lead to ill-defined backward no-branching probabilities.

3.2 Event generation techniques

Event generation according to the above defined merging procedure proceeds along the following lines:

- Relevant multi-jet cross sections for the process under consideration are calculated with the phase-space restriction $Q > Q_{\rm cut}$. Strong couplings are computed such that they give an overestimate, which can later on be reweighted. PDF's are evaluated at the scale set by the core process.

1. Events are generated according to the above defined cross sections with kinematics determined by the respective matrix elements.

2. Matrix elements are interpreted in the large N_C limit according to methods discussed in Sec. 2.5. The most probable shower history of the final state is determined through backward clustering, cf. Sec. 2.3. The clustering is guided by information from the matrix element, which restricts the available shower histories to those, which correspond to a Feynman diagram.

3. The event is accepted or rejected according to a kinematics-dependent weight, which corresponds to evaluating strong couplings in the shower scheme and computing the no-branching probability, Eq. (18), for each parton, internal or external, in the clustered matrix element.

4. The parton-shower evolution is started with suitably defined scales for intermediate and final-state particles. During showering, any emission harder than Q_{cut} is vetoed. Intermediate partons undergo a truncated shower evolution, cf. Sec. 2.4

This algorithm has the apparent drawback, that the no-emission probability Eq. (18) must be computed before the parton-shower evolution starts. Ideally, however, it should result as a direct consequence of parton-shower branchings producing hard partons. Such splittings would obviously fall in the realm of matrix elements and are thus forbidden inside the parton shower. To obtain a corresponding prescription, the above algorithm is slightly reformulated.

Firstly, the logarithmic derivative of the no-branching probabilities $\mathcal{P}_{\text{no},\,a}^{(B)\,\text{ME/PS}}$ is defined as

$$\mathcal{I}_a^{(B)\,\text{ME/PS}}(z,\bar{t}) = \int_z^{\zeta_{\max}} \frac{\mathrm{d}\zeta}{\zeta} \sum_{b=q,g} \mathcal{K}_{ba}^{\text{ME/PS}}(\zeta,\bar{t})\, \frac{g_b(z/\zeta,\bar{t})}{g_a(z,\bar{t})}\,. \tag{19}$$

From Eq. (7) one then obtains the full branching probability in terms of $\mathcal{I}^{(B)\,\text{ME/PS}}$,

$$\mathcal{P}_{\text{branch},\,a}^{(B)\,\text{ME}\oplus\text{PS}}(z,t,t') = \Big[\, \mathcal{I}_a^{(B)\,\text{ME}}(z,t) + \mathcal{I}_a^{(B)\,\text{PS}}(z,t) \,\Big] \, \exp\left\{ -\int_t^{t'} \frac{\mathrm{d}\bar{t}}{\bar{t}}\, \mathcal{I}_a^{(B)}(z,\bar{t}) \right\}\,, \tag{20}$$

where $\mathcal{I}_a^{(B)} = \mathcal{I}_a^{(B)\,\text{ME}} + \mathcal{I}_a^{(B)\,\text{PS}}$. Equation (20) corresponds to generating an ordering parameter t in unconstrained shower evolution, i.e. without the restriction to $Q < Q_{\text{cut}}$. The first term in the square bracket is however given by hard matrix elements through Eq. (2). In order not to double count this contribution, corresponding branchings must lead to rejection of the entire event. This modifies the respective cross section by

$$\sigma \to \sigma \cdot \mathcal{P}_{\text{no},\,a}^{(B)\,\text{ME}}(t,t')\,. \tag{21}$$

Due to this event rejection, the remaining branching probability for accepted parton-shower steps is given by (cf. the description of the veto algorithm, for example in [40])

$$\begin{aligned} \mathcal{P}_{\text{branch},\,a}^{(B)\,\text{PS}}(z,t,t') = \mathcal{I}_a^{(B)\,\text{PS}}(z,t)\, \exp\left\{ -\int_t^{t'} \frac{\mathrm{d}\bar{t}}{\bar{t}}\, \mathcal{I}_a^{(B)}(z,\bar{t}) \right\} \\ \times \exp\left\{ \int_t^{t'} \frac{\mathrm{d}\bar{t}}{\bar{t}}\, \Big[\mathcal{I}_a^{(B)}(z,\bar{t}) - \mathcal{I}_a^{(B)\text{PS}}(z,\bar{t}) \Big] \right\}\,, \end{aligned} \tag{22}$$

which yields the vetoed shower algorithm described by Eq. (16), but with $\tilde{g} = g$, as required.
We therefore obtain the modified rules

3. The event is accepted or rejected according to a kinematics dependent weight, which corresponds to evaluating strong couplings in the parton-shower scheme.

4. The parton-shower evolution is started with suitably defined scales for intermediate and final-state particles. Intermediate partons undergo truncated evolution. During showering, any emission harder than Q_{cut} leads to the rejection of the event.

Note that in prinicple these two steps could be combined through evaluating the strong couplings during the shower evolution.

The relation of this algorithm with other merging prescriptions is discussed in Appendix A.

3.3 Highest multiplicity treatment

An apparent problem of the merging algorithm outlined so far is that only a limited final-state multiplicity $N \leq N_{\max}$ can be generated through full matrix elements. Hence the matrix elements will not produce jet

multiplicities $N > N_{\max}$ that are in principle possible. Hence, the parton shower must account for missing emissions above Q_{cut} at large N. This is explained in detail in the following.

Assume a case where $N = N_{\max}$ emissions in the matrix-element domain have been accounted for by the matrix element and have been generated through the above defined algorithm. This means that up to this point, i.e. up to the scale where the last matrix-element emission can be resolved in terms of the parton-shower evolution parameter, the branching probability, Eq. (20) has been employed, as it should be. Beyond this point, no further emission can be generated through matrix elements, and the branching probability becomes

$$\mathcal{P}^{(B)\,\mathrm{ME}\oplus\mathrm{PS}}_{\mathrm{branch},\,a}(z,t,t') \to \mathcal{P}^{(B)\,\mathrm{ME}}_{\mathrm{no},\,a}(t,t')\,\frac{\partial\mathcal{P}^{(B)\,\mathrm{PS}}_{\mathrm{no},\,a}(z,t,t')}{\partial\log(t/\mu^2)}\,. \tag{23}$$

Relation (23) would obviously violate factorisation, because of missing terms, corresponding to the integrated kernel from the matrix-element domain.

This problem can be circumvented by implementing the standard parton-shower evolution beyond the last matrix-element emission.[4] It guarantees that the parton shower respects the description of hard radiation throughout the regime where matrix elements are applicable, while still filling the remaining phase space.

This prescription is referred to as the highest multiplicity treatment and has been suggested in a similar form in [8]. In virtuality ordered DGLAP evolution, it approximately corresponds to setting a local veto scale $Q_{\text{cut}} \to Q_{\min}$ if $N = N_{\max}$, where $Q_{\min}$ is the minimum jet criterion found during backward clustering.

3.4 Sources of uncertainties

The proposed merging algorithm combines two essentially different approaches to perturbative QCD. Any simulation based on it therefore contains a number of sources of theoretical uncertainties. They can be separated into two categories, merging-related and non-merging-related uncertainties. The latter would occur in standard perturbative approaches as well, when using only hard matrix elements or applying only parton showers. The merging-related uncertainties are instead specific for the combined application of matrix elements and showers and arise from the following:

- The specific choice of the jet criterion.
  Since the jet criterion separates matrix-element and parton-shower domain, a variation of its precise definition can enhance or reduce the contribution of the hard matrix elements in certain regions of phase space.

- The value of the phase-space separation cut, Q_{cut}.
  As for the jet criterion itself, the precise value of the separation cut enhances or reduces the amount of phase space which is described by hard matrix elements and can therefore lead to variation of the results.

- The maximum number of jets simulated by hard matrix elements, $N_{\max}$.
  This parameter limits the number of hard partons up to which correlations can be expected to be correctly described at tree-level.

Other uncertainties are related to the perturbative calculations carried out in the matrix-element and parton-shower simulation itself. They include:

- Scale uncertainties from matrix elements.
  They arise due to the particular choice of factorisation and renormalisation scale of the leading-order process.

- Scale uncertainties from parton showers.
  They arise due to the particular choice of coupling scales within the evolution.

- Uncertainties from the parton density functions employed.
  Parton density functions not only enter the cross section calculation when considering hadronic initial

[4] The term "beyond" refers to the ordering parameter t. Note that the respective scale is set globally for the event, because the matrix element connects all parton-shower evolutions.

states but also appear in the calculation of the branching probabilities for the initial-state parton shower, for a detailed discussion of the latter see e.g. [41].

- Uncertainties due to the choice of the leading-order process.
  These uncertainties arise in processes which potentially contain many additional jets with shower evolution parameters above the factorisation scale of the leading-order process. Corresponding details are discussed in Appendix B.

In this publication we focus on a study of the pure merging-related uncertainties. We will, however not vary the jet criterion, but rather employ what we identify as the optimal choice for the merging, see Sec. 4.

4 The jet criterion

An important aspect in the QCD evolution equations, Eq. (1), is that QCD branchings are logarithmically enhanced at small values of the evolution parameter t and/or at logarithmically large values of the evolution kernels $\mathcal{K}$. This is the manifestation of the singular infrared behaviour of QCD amplitudes in the respective regions of phase space. In perturbative calculations employing hard matrix elements these regions therefore must be regularised. This is typically achieved by identifying parton samples or individual partons with jets and demanding the jets to be sufficiently isolated. Algorithms defining jets are, for example, the Durham k_T-algorithm [10] and the longitudinally invariant k_T algorithms for deep inelastic scattering and hadron-hadron collisions [11]. Extensions of those algorithms to include jet flavour have been presented in [42]. Their respective measures are often used as a variable in which phase-space separation is defined for matrix element - parton shower merging, cf. [6, 7].

We propose a similar criterion here. However, in contrast to jet measures like the ones above, which can be applied to experimentally observable final states and which yield experimentally well defined jets, this criterion is designed and applied on purely theoretical grounds. It is based on the behaviour of QCD amplitudes at the next-to-leading order and employs flavour and colour information of the respective partons. For our purposes this criterion proved to be advantageous over standard k_T algorithms as it correctly identifies individual infrared enhanced QCD branchings.

4.1 Definition

Consider two partons i and j, which can, in terms of flavour and colour, originate from a common mother parton (the splitter) $\widetilde{ij}$. The following jet criterion is then proposed

$$Q_{ij}^2 \;=\; 2\,p_i p_j \,\min_{k\neq i,j}\, \frac{2}{C_{i,j}^k + C_{j,i}^k}\;, \tag{24}$$

where for final state partons i and j

$$C_{i,j}^k \;=\; \begin{cases} \dfrac{p_i p_k}{(p_i+p_k)p_j} - \dfrac{m_i^2}{2\,p_i p_j} & \text{if } j = g \\ 1 & \text{else} \end{cases}\;. \tag{25}$$

For initial state partons a, we consider the splitting process $a \to (aj)\; j$. With the momentum of the combined particle (aj) given by $p_{aj} = p_a - p_j$, we define

$$C_{a,j}^k \;=\; C_{(aj),\,j}^k\;. \tag{26}$$

The minimum in Eq. (24) is over all possible colour partners k of the combined parton $\widetilde{ij}$, which can be thought of to act as spectators in the splitting process.

In the following, it is shown that this jet criterion indeed correctly identifies soft and collinear parton splittings in QCD matrix elements and is thus suited to separate the matrix-element from the parton-shower domain in the merging.

4.2 Soft limit

If the energy of a single gluon j tends to zero in any fixed direction q, described through $p_j = \lambda q$, $\lambda \to 0$, the above jet criterion behaves as

$$\frac{1}{Q_{ij}^2} \to \frac{1}{2\,\lambda^2}\;\frac{1}{2\,p_i\,q}\;\max_{k\neq i,j}\left[\frac{p_i p_k}{(p_i+p_k)\,q} - \frac{m_i^2}{2\,p_i q}\right]\,. \tag{27}$$

The corresponding singularity of the matrix element is thus correctly identified, cf. [43].

4.3 Quasi-collinear limit for final-state splittings

Consider two final-state partons i and j and an arbitrary spectator-parton k. Let $p_{ij} = p_i + p_j$ and let the light-like helper vectors l and n be defined by

$$\begin{aligned} p_{ij} &= l + \alpha_{ij}\,n\,, \\ p_k &= n + \alpha_k\,l\,. \end{aligned} \tag{28}$$

This system has the solution

$$l = \frac{1}{1-\alpha_{ij}\alpha_k}\,(p_{ij} - \alpha_{ij}\,p_k)\,, \qquad\qquad n = \frac{1}{1-\alpha_{ij}\alpha_k}\,(p_k - \alpha_k\,p_{ij})\,, \tag{29}$$

where $\alpha_{ij} = p_{ij}^2/\gamma$, $\alpha_k = p_k^2/\gamma$ and $\gamma = 2\,ln = p_{ij}p_k + \sqrt{(p_{ij}p_k)^2 - p_{ij}^2 p_k^2}$, cf. [44]. The momenta p_i and p_j can now be expressed in terms of l, n and a transverse component, $k_\perp$.

$$\begin{aligned} p_i^\mu &= z\,l^\mu + \frac{m_i^2 + \mathrm{k}_\perp^2}{z}\,\frac{n^\mu}{2\,ln} + k_\perp^\mu\,, \\ p_j^\mu &= (1-z)\,l^\mu + \frac{m_j^2 + \mathrm{k}_\perp^2}{1-z}\,\frac{n^\mu}{2\,ln} - k_\perp^\mu\,. \end{aligned} \tag{30}$$

A relation for p_{ij}^2 is immediately obtained,

$$p_{ij}^2 - m_i^2 - m_j^2 = \frac{\mathrm{k}_\perp^2}{z(1-z)} - \frac{1-z}{z}\,m_i^2 - \frac{z}{1-z}\,m_j^2\,. \tag{31}$$

Taking the quasi-collinear limit amounts to the simultaneous rescaling [45]

$$\mathrm{k}_\perp \to \lambda \mathrm{k}_\perp\,, \qquad m_i \to \lambda m_i\,, \qquad m_j \to \lambda m_j\,, \qquad m_{ij} \to \lambda m_{ij}\,. \tag{32}$$

Then, $2\,p_i p_j \to \lambda^2(p_{ij}^2 - m_i^2 - m_j^2)$ and, independent of k,

$$\frac{1}{Q_{ij}^2} \to \frac{1}{2\,\lambda^2}\;\frac{1}{p_{ij}^2 - m_i^2 - m_j^2}\,\left(\tilde{C}_{i,j} + \tilde{C}_{j,i}\right)\,. \tag{33}$$

Here,

$$\tilde{C}_{i,j} = \begin{cases} \dfrac{z}{1-z} - \dfrac{m_i^2}{2\,p_i p_j} & \text{if } j = g \\[2ex] 1 & \text{else} \end{cases}\,. \tag{34}$$

Equation (34) corresponds to the leading term of the massive Altarelli-Parisi splitting function for $z \to 1$ [45]. The corresponding term for $z \to 0$ (if present) is generated by $\tilde{C}_{j,i}$.

4.4 Quasi-collinear limit for initial-state splittings

Now consider the initial-state parton a, the final-state parton j and an arbitrary spectator parton k. Let $p_{aj} = p_a - p_j$, and let the light-like helper vectors l and n be defined by

$$\begin{aligned} p_a &= l + \alpha_a\,n\,, \\ p_k &= n + \alpha_k\,l\,. \end{aligned} \tag{35}$$

Then l and n are found as before, Eq. (29). The momenta p_{aj} and p_j are decomposed as follows

$$\begin{aligned} p_{aj}^\mu &= z\, l^\mu + \frac{p_{aj}^2 + \mathrm{k}_\perp^2}{z}\, \frac{n^\mu}{2\, ln} + k_\perp^\mu \,, \\ p_j^\mu &= (1-z)\, l^\mu + \frac{m_j^2 + \mathrm{k}_\perp^2}{1-z}\, \frac{n^\mu}{2\, ln} - k_\perp^\mu \,. \end{aligned} \tag{36}$$

Taking the quasi-collinear limit yields $2\, p_a p_j \to \lambda^2 \left| p_{aj}^2 - m_a^2 - m_j^2 \right|$ such that, independent of k

$$\frac{1}{Q_{aj}^2} \to \frac{1}{2\,\lambda^2}\, \frac{1}{\left| p_{aj}^2 - m_a^2 - m_j^2 \right|} \left(\tilde{C}_{(aj),j} + \tilde{C}_{j,(aj)} \right) \,, \tag{37}$$

where $\tilde{C}_{(aj),j}$ is given by Eq. (34).

5 Monte Carlo programs for the study

In this section, we present the Monte Carlo programs that have been employed in the actual implementation of the merging algorithm outlined in Sec. 3. As discussed in Sec. 3.4, the foremost aim for this publication is to study the specific systematics of the proposed merging procedure.

For tree-level amplitudes there exists now a full wealth of publicly available programs that accomplish the evaluation of (nearly) arbitrary multi-parton processes within the Standard Model and some of its most prominent extensions. These highly automated tools, called matrix-element or parton-level generators, either rely on the translation of Feynman diagrams into helicity amplitudes [46] or make use of recursive relations to obtain compact expressions for the amplitudes [47, 48]. In addition to the pure matrix elements they provide suitable phase-space integrators to permit the evaluation of cross sections and the generation of actual parton level events. In the context of this work we employ two different matrix-element generators, the well-established program AMEGIC++ [49] and the recently released code COMIX [50]. Both codes are integrated into the multi-purpose event generator SHERPA [12] that steers the event generation and hosts various interfaces to parton showers, hadronisation routines, analysis methods and the like. The two programs share the capability to simulate arbitrary Standard Model processes at tree-level. Their main difference lies in the treatment of colour, providing us with a handle on the related systematic uncertainties. We will briefly repeat the main features of both approaches in Secs 5.1 and Sec. 5.2, respectively.

Contrary to the matrix-element generators, there exist only few publicly available implementations of parton-shower algorithms. There is the class of PYTHIA-like parton showers using either virtuality [51] or transverse momentum [52] as evolution variable. On the other hand there are implementations of angular-ordered parton showers in the HERWIG and HERWIG++ generators, see for instance [53] and references therein. A shower implementation based on colour dipoles became available with the release of the ARIADNE [30] program. For our study we employ the parton-shower model presented in [35]. This approach relies on Catani–Seymour dipole factorisation of QCD amplitudes and organises subsequent parton emissions in terms of an invariant transverse momentum. Aspects of this parton-shower model relevant for combining it with matrix element calculations will be discussed in Sec. 5.3.

5.1 The matrix-element generator AMEGIC++

AMEGIC++ [49] is a tree-level matrix-element generator, based on Feynman diagrams. To evaluate the single amplitudes of a given process, the helicity methods introduced in [54] are employed. Diagrams are constructed and sorted according to their respective colour structure. A colour matrix for the full squared matrix element is computed using standard methods. Each single Feynman diagram is then decomposed into basic building blocks in the helicity formalism. It is important to note in this context, that information about the diagrams can be accessed during the runtime of SHERPA and therefore clustering sequences within the merging formalism can be chosen according to the propagator structure of the contributing graphs. Thus, unphysical combinations can be prevented. The fact that the treatment of colour in AMEGIC++ essentially follows textbook methods is however rather problematic in the context of a matrix element - parton shower merging. Since it is not possible within the code to access colour-ordered amplitudes, a "most probable" colour structure must be selected on kinematic grounds. This then yields the nodal values to define starting conditions for parton showers.

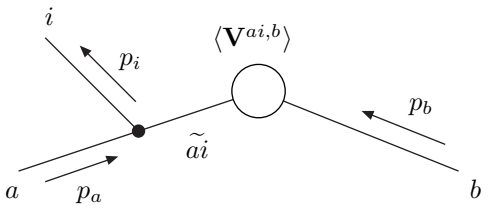

Figure 1: Schematic view of the splitting of an initial-state parton with an initial-state spectator. The blob denotes the n-parton matrix element. Incoming and outgoing lines label initial- and final-state partons, respectively.

5.2 The matrix-element generator COMIX

COMIX is based on an extension of the colour-dressed Berends-Giele recursive relations for QCD amplitudes [37] to the full Standard Model. These relations are essentially equivalent to the Dyson-Schwinger recursion employed for instance in [47]. For any recursive calculation, the growth in computational complexity of the algorithm solely depends on the number of external legs at elementary vertices of the theory. Thus within COMIX any four-particle vertex of the Standard Model is decomposed into three-particle vertices [50]. This leads to an improved performance for large-multiplicity matrix elements, compared to AMEGIC++. The summation (averaging) over colours in QCD and QCD-associated processes is performed in a Monte Carlo fashion and colour-ordered amplitudes can therefore be computed. Following the reasoning of [37], the colour-flow basis is employed throughout the code. As discussed in [36] and Sec. 2.5, this yields a certain correspondence between the large-N_C limit employed in parton-shower simulations and full QCD results. COMIX allows to access current information during event generation, such that, like for AMEGIC++, unphysical clusterings in the matrix element - parton shower merging can be prevented. The corresponding algorithm respects the actual colour assignment of external states, cf. Sec. 2.5.

5.3 The Catani-Seymour subtraction based shower generator

In this work we employ a shower approach based on Catani-Seymour (CS) dipole factorisation, which will be denoted by CSS [35]. The model was originally proposed in [33] and worked out and implemented in [34, 35]. It relies on the factorisation of real-emission matrix elements in the CS subtraction framework [55, 43]. There exist four general types of CS dipole terms that capture the complete infrared singularity structure of next-to-leading order QCD amplitudes, namely splitting initial-state particles accompanied by an initial or final-state parton as spectator, or branching final-state lines associated with either another final-state leg or an incoming parton as spectator. In the large-N_C limit, the splitter and spectator partons are always adjacent in colour space. The dipole functions for the various cases, taken in four dimensions and averaged over spins, are used as shower splitting kernels. Their infrared singularities are regularised through a finite cutoff parameter of order Λ_{QCD}, the shower stopping scale. Consider, as an example, the case of a splitting initial-initial CS dipole, cf. Fig. 0. Following the nomenclature used in [35], we can then, in analogy to Eq. (3), identify

$$\mathcal{K} \to \mathcal{K}_{a(ai)}(x_{i,ab}, \mathbf{k}_\perp^2) = \frac{\alpha_s(\mathbf{k}_\perp^2/4)}{2\pi} \langle \mathbf{V}^{ai,b}(x_{i,ab}) \rangle \,. \tag{38}$$

The actual splitting is parametrised in terms of the Lorentz invariants

$$x_{i,ab} = \frac{p_a p_b - p_i p_a - p_i p_b}{p_a p_b} \quad \text{and} \quad \mathbf{k}_\perp^2 = 2\tilde{p}_{ai} p_b \, \tilde{v}_i \, \frac{1 - x_{i,ab}}{x_{i,ab}} \,, \quad \text{where} \quad \tilde{v}_i = \frac{p_i p_a}{p_a p_b} \,. \tag{39}$$

The invariant transverse momentum $\mathbf{k}_\perp^2$ acts as ordering parameter for initial-state splittings in the actual

shower algorithm[5]. The three other CS dipole configurations are treated following a similar reasoning as just presented for initial-initial dipoles. Corresponding Sudakov form factors for all branching types, taking into account finite parton masses and strictly relying on Lorentz-invariant variables, have been derived in [35].

All branchings in the CSS formalism implement exact four-momentum conservation and the particles after an evolution step are kept on their mass-shell. The phase-space maps from an n to an $n+1$ particle final state are exact and cover the whole phase space. With the only exception of splitting CS initial-initial dipoles, the momentum recoil of a certain branching is compensated locally by the assigned spectator parton. For splitting initial-initial dipoles the spectator momentum is kept fixed, consequently the recoil is taken by the entire ensemble of final-state particles.

The recoil strategy in this shower scheme, in conjunction with the strictly Lorentz-invariant evolution parameters and splitting variables, eventually allows the construction of a clustering algorithm along the lines of Sec. 2.3 and a truncated shower, cf. Sec. 2.4. Both are necessary ingredients of the merging formalism presented in this publication. The main difference with respect to other shower formulations, in this respect, lies in the fact that the recoil partner of a splitting parton is always a single external line of the current initial or final state[6]. If this is not guaranteed, the implementation of a clustering algorithm as proposed in Sec. 2.3 becomes cumbersome.

Another apparent advantage of this parton-shower model is that it inherently respects QCD soft colour coherence. By construction in Catani–Seymour factorisation the eikonal factor associated to soft gluon emission off a colour dipole, used to derive the angular ordering constrained in conventional parton showers, is exactly mapped onto two CS dipoles. These two dipoles just differ by the role of emitter and spectator, see [55]. Accordingly, the shower algorithm based on Catani–Seymour dipole factorisation captures the essential contributions of the next-to-leading logarithmic corrections, beyond the naive resummation of the leading soft and collinear enhanced terms. For a further discussion on related issues, the reader is referred to [56]. For a comparison of the shower model at hand with experimental data sensitive to the correct modelling of soft-gluon emission see [35].

A potential shortcoming of the shower algorithm based on Catani–Seymour dipole terms is that certain dipole functions connecting the initial and final state may assume negative values in some non-singular phase-space regions. This prohibits their naive interpretation in terms of splitting probabilities, cf. [34, 35]. This problem, however, can now partially be cured once shower emissions are corrected by exact matrix-element calculations that provide the appropriate description of QCD amplitudes in these regions of phase space.

6 Results

In this section we present results generated with the Monte Carlo programs described previously. The main purpose of this brief study is to quantify some systematic uncertainties introduced by the proposed merging algorithm, cf. Sec. 3.4. Therefore, three main questions are considered:

- Are observables sufficiently independent on the "unphysical" phase-space separation criterion Q_{cut}?
- How does the description of additional jets by the matrix element improve the prediction of experimentally relevant quantities?
- As a combination of the above: Is the total cross section as predicted by the lowest order unaltered to the required accuracy?

Using the two matrix-element generators AMEGIC++ and COMIX, together with the shower generator CSS in the framework of SHERPA, we also have the opportunity to study effects of using different methods to assign colours in the large N_C limit.

In the first part of the section we focus on e^+e^- annihilation into hadrons. This setup allows to study the algorithm in pure final-state QCD evolution. The influence of different colour-assignment strategies is

[5] Note that the actual form of $\mathbf{k}_\perp^2$ is different to the one in [35], allowing to suppress an additional Jacobian factor obtained ibidem. We employ this same quantity as ordering parameter for both dipole types with initial-state splitters, whereas the choice for dipoles with final-state splitter remains unchanged with respect to [35], namely the invariant transverse momentum squared of the two splitting products.

[6] This is also true for initial-initial dipoles, since the shower formulation is invariant and the reference frame can be chosen as the centre-of-mass frame of the process before the splitting.

		$N_{\max}$				
		0	1	2	3	4
$\log_{10}(Q^2_{\text{cut}}/s)$	-1.25	40.17(1)	39.65(3)	39.66(3)	39.66(3)	39.67(3)
	-1.75		39.38(5)	39.29(6)	39.13(5)	39.13(5)
	-2.25		39.27(8)	38.35(9)	37.89(11)	37.60(10)

Table 1: Total cross sections for $e^+e^- \to jj$ at $\sqrt{s} = 91.25$GeV in [nb] and their dependence on the separation criterion and the maximum number of additional jets produced in the matrix element.

investigated in this configuration. The second part deals with Drell-Yan lepton pair production in hadronic interactions, which is used to validate the algorithm in combined initial and final-state evolution.

Jets at parton level include the quark flavours d, u, s, c and b as well as gluons. The PDF set employed is CTEQ 6L [57], which defines the corresponding α_s parametrisation in hadron collisions. All other generator parameters are left at the default values of the Monte Carlo programs, since none of them has any impact on the QCD predictions.

Hadron-level results are produced using SHERPA with the fragmentation module AHADIC++ [58], the hadron and τ decay package HADRONS++ [59] and a simulation of extra QED radiation through PHOTONS++ [60]. The hadronisation phases remain untuned so far and will be tuned to data at a later stage with the help of the PROFESSOR package [61]. The RIVET package [62] is employed for analysis and comparison to data.

6.1 QCD jet production in e^+e^- collisions

To compare with LEP data from the Run I period, a setup with cms energy $\sqrt{s} = 91.25\,\text{GeV}$ is chosen and a merged sample of $e^+e^- \to (N+2)$ jets is produced at parton level, with N denoting the number of additional jets in the matrix element and $N \leq N_{\max}$. We vary $N_{\max}$ between zero, i.e. no merging at all, and four.

Total cross sections and jet rates

Firstly, we present a comparison of total cross sections predicted by the merging algorithm for various values of the separation criterion Q_{cut} and the maximum jet multiplicity $N_{\max}$. Table 0 shows only minor deviations, i.e. up to 6.4%, between the leading-order cross section and predictions for the merged samples. We can thus confirm that the proposed merging approach preserves the cross section of the leading-order process.

Figure 1 shows integrated rates of jets determined with the Durham k_T-algorithm [10] as a function of the analysis cut y_{cut}. As well as giving a fine-grained insight into the number of jets to be expected for a given analysis cut, they also provide the assurance that the merging algorithm gives accurate predictions for the perturbative region in which it operates. Jets found with cuts below the shower regime, $y_{\text{cut}} \approx 10^{-3.5}$, are influenced by hadronisation effects, and are thus not relevant within the scope of this paper. Even scales slightly above might be populated e.g. by decays of heavy mesons. Monte Carlo results have been produced using a merged sample for up to four additional jets in the final state, generated with COMIX and showered with CSS with a merging cut $\log_{10}(Q^2_{\text{cut}}/s) = 2.25$.

Differential distributions

Due to the close correspondence between the jet criterion, Eq. (24), and the Durham measure for jets in e^+e^- collisions (cf. also Appendix A.1), differential jet rates are particularly suited to assess merging systematics. The rate $y_{n\,n+1}$ shows the jet measure at which $n+1$ jets are clustered into n jets according to the Durham k_T-algorithm. The phase-space separation cut Q_{cut} manifests itself as a narrow transition region between matrix-element and parton-shower domain around $y_{n\,n+1} \approx Q^2_{\text{cut}}/s$. Merging systematics can thus be inferred from deviations between samples with different phase-space separation cuts in this region.

Figure 2 shows the differential jet rates for a merged sample of up to four additional jets from the matrix element, generated with COMIX and showered with the CSS. The merging cuts, which have been used, are

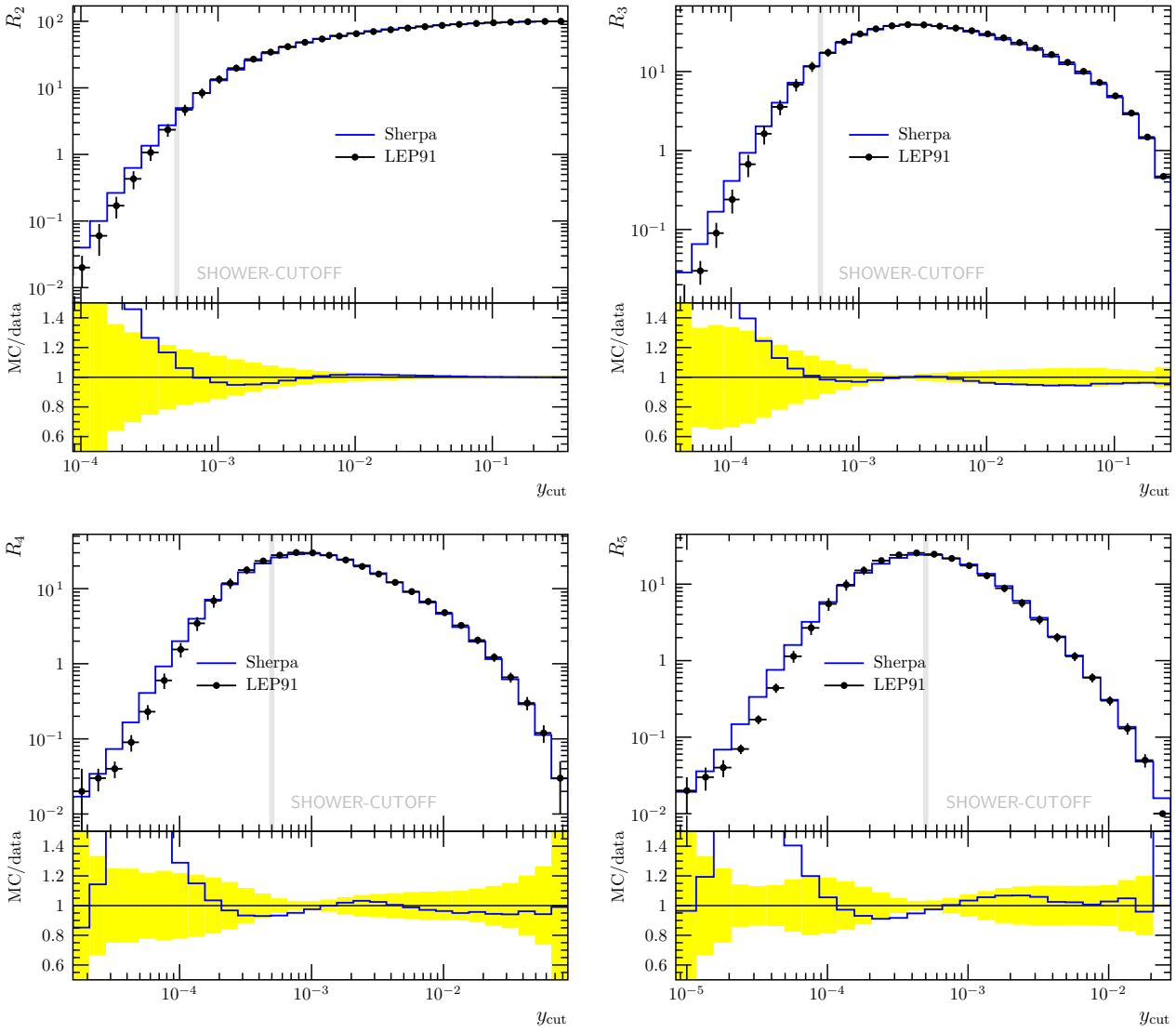

Figure 2: Integrated jet rates compared to data from LEP [63].

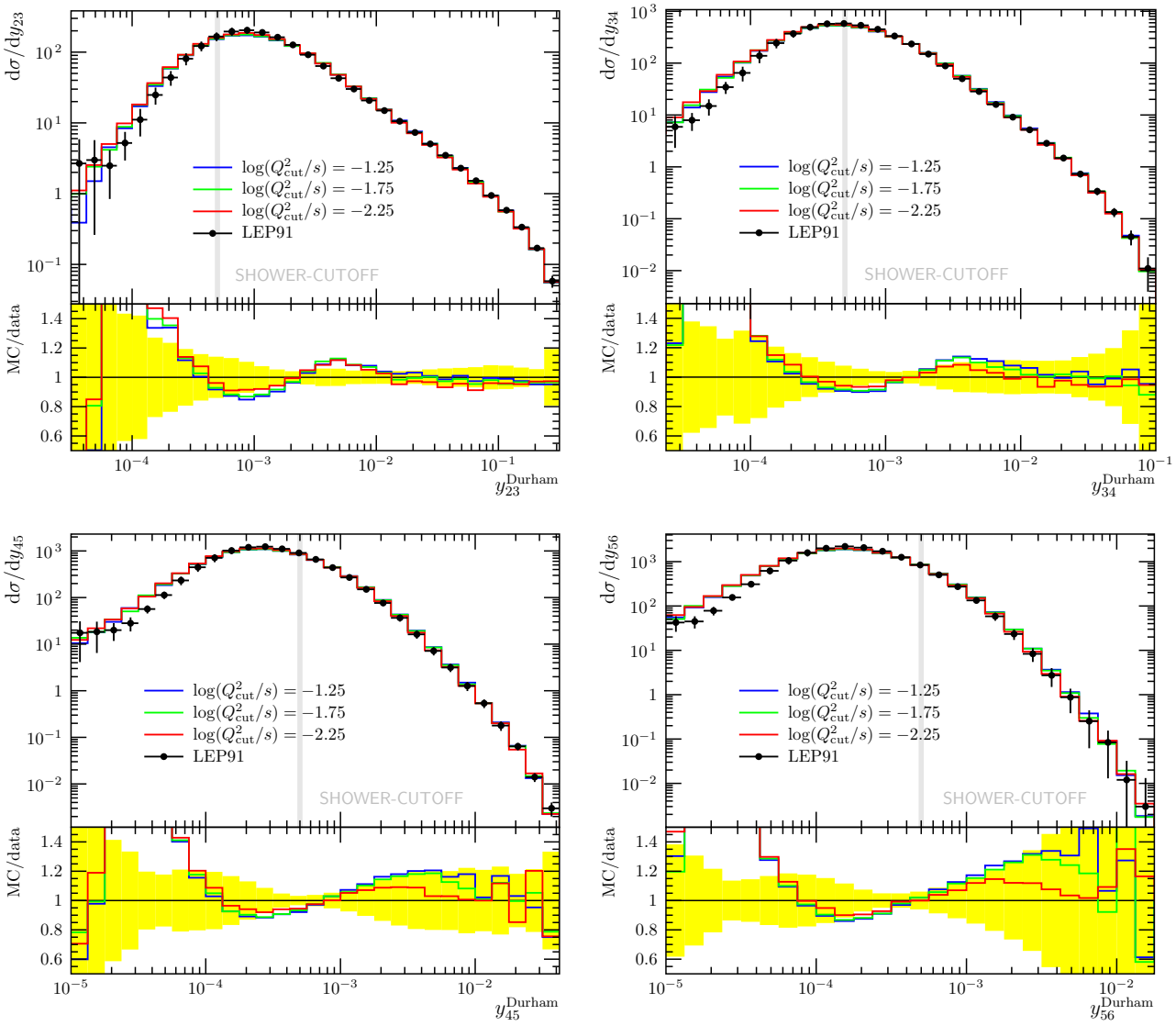

Figure 3: Differential jet rates for three different merging cuts compared to data from LEP [63].

$\log_{10}(Q^2_{\text{cut}}/s) = -2.25$, $\log_{10}(Q^2_{\text{cut}}/s) = -1.75$ and $\log_{10}(Q^2_{\text{cut}}/s) = -1.25$. We observe only tiny deviations between the predictions of the various samples.

Event-shape observables like thrust, sphericity and the energy-energy correlation are presented in Fig. 3. Details about their definition and the corresponding data from experiment can be found in [64]. These observables are well described by an appropriately tuned pure parton-shower setup already, and no matrix element improvement is therefore necessary. On the other hand, the comparison between the pure parton-shower sample and merged samples constitutes an important consistency check. We find very good agreement of the respective predictions.

Colour assignment at large N_C

To investigate the influence of different strategies to handle colour in matrix-element generation and merging, we compare the generators COMIX and AMEGIC++. COMIX is run in two different modes, corresponding to the two strategies of assigning colours in the large-N_C limit explained in Sec. 2.5. More precisely we denote by

'partial σ'
: selection of a colour assignment according to the proportionate squared partial amplitude corresponding

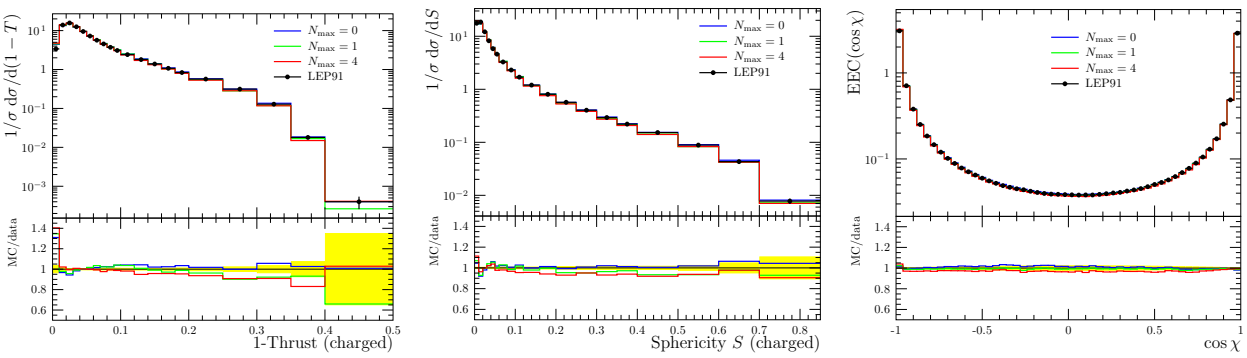

Figure 4: Event-shape observables compared to data from LEP [64].

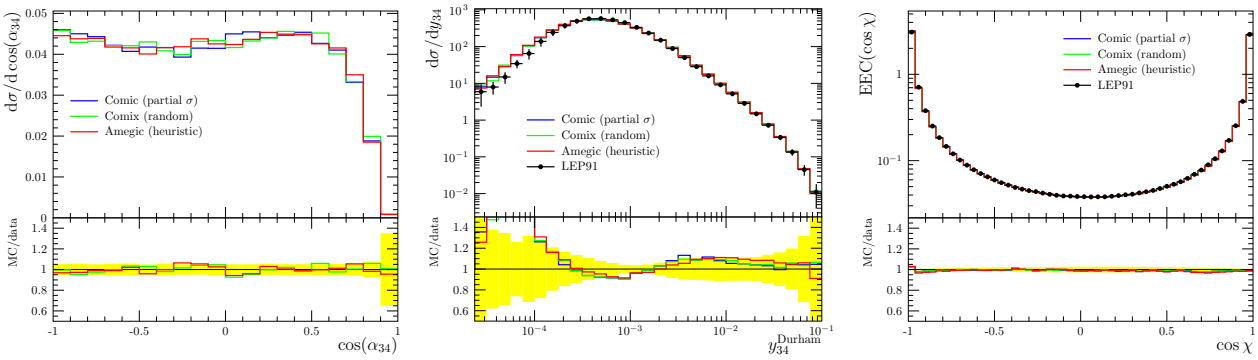

Figure 5: Different approaches to colour treatment and their effects on y_{34}, α_{34} and EEC. Data are taken from [63, 64].

to this large-N_C configuration (i.e. the *first* method presented in Sec. 2.5)

'random'
selection of a colour assignment according to the ratio of the squared partial amplitude corresponding to this large-N_C configuration and the total amplitude squared (i.e. the *second* method presented in Sec. 2.5)

'heuristic'
heuristic colour assignment to the previously colour-summed amplitudes in AMEGIC++ (cf. Sec 5.1)

Figure 4 shows respective predictions for the three different choices. Since the first configuration where different colour assignments could take effect arises in four-jet events, the selected observables are the $3 \to 4$ jet rate and the angle α_{34} between the two softest jets, selected on an event-by-event basis. Furthermore, the energy-energy correlation typically shows sensitivity to the connection of the hadronisation phase to the parton-shower output, and could thus depend on the colour setting as well. We observe no significant differences between the various options. This encourages to proceed with even the heuristic method, which enables us to employ the merging technique with various kinds of matrix element generators, including those which do not allow a projection onto the large-N_C limit.

6.2 Drell-Yan lepton production in $p\bar{p}$ collisions

The scope of this section is to validate the proposed merging algorithm in collisions with hadronic initial states. One of the simplest processes in this setup is Drell-Yan lepton pair production. It constitutes an important testing ground to validate the applicability of the proposed jet criterion and the interplay of the merging algorithm with the PDF's.

Event generation has been set up for $p\bar{p}$-collisions at a centre-of-mass energy of $\sqrt{s} = 1960\,\mathrm{GeV}$. For the hard process a merged sample of $p\bar{p} \to e^+e^- + N\,\mathrm{jets}$ has been produced, where $N \leq N_{\max}$ with

		$N_{\max}$						
		0	1	2	3	4	5	6
Q_{cut}	20 GeV	192.6(1)	191.0(3)	190.5(4)	189.0(5)	189.4(7)	188.2(8)	189.9(10)
	30 GeV		192.3(2)	192.7(2)	192.6(3)	192.9(3)	192.7(3)	193.2(3)
	45 GeV		193.6(1)	194.4(1)	194.3(1)	194.4(1)	194.6(2)	194.4(1)

Table 2: Total cross sections [pb] in $p\bar{p} \to e^+e^- + \text{jets}$ at $\sqrt{s} = 1960\,\text{GeV}$ and their dependence on the merging cut.

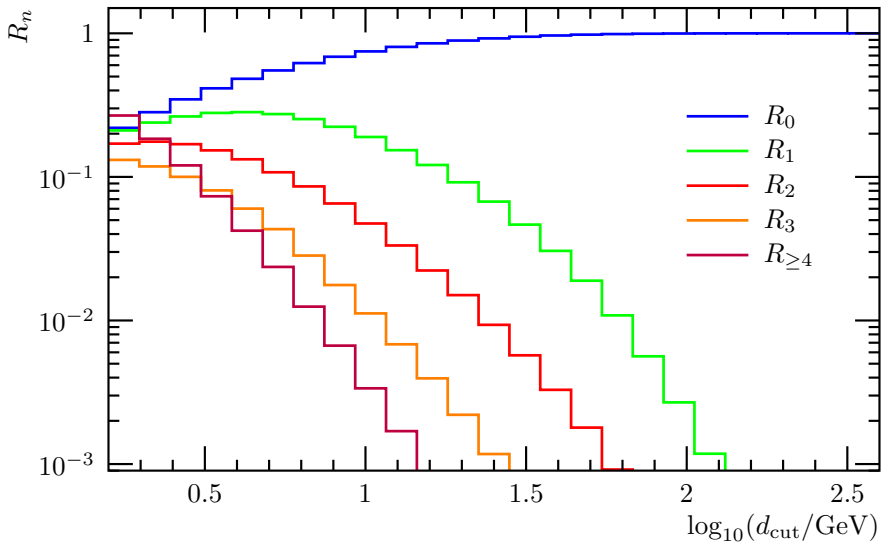

Figure 6: Integrated jet rates as a function of the analysis cut.

$0 \leq N_{\max} \leq 6$. In addition, a cut of $66\,\text{GeV} < m_{e^+e^-} < 116\,\text{GeV}$ has been applied at the matrix-element level. The factorisation scale has been chosen as $m^2_{e^+e^-}$. Note that the transverse mass squared of the lepton pair, $m^2_{T,e^+e^-} = m^2_{e^+e^-} + \text{k}^2_{\perp,e^+e^-}$, which is often selected as the factorisation scale in other merging approaches, is not a proper choice for the proposed algorithm. It is non-continuous with respect to the transverse momentum, $\text{k}^2_{\perp,e^+e^-}$, because the leading-order configuration is generated with $\text{k}^2_{\perp,e^+e^-} = 0$ and the minimum transverse momentum of events with one additional jet is limited by the phase-space separation cut.

Total cross sections and jet rates

Again, we first present a comparison of total cross sections predicted by the merging algorithm for various values of the separation criterion Q_{cut} and the maximum jet multiplicity $N_{\max}$. Table 1 shows the respective results. Differences range up to 2.3%, between the leading-order cross section and predictions for the merged samples. Usually, the systematic uncertainties in hadronic collisions are larger than in e^+e^- reactions, partly due to PDF uncertainties and partly because the shower evolution generated by the CSS and the resummation used to compute the PDF's are not entirely compatible, cf. [55]. This effect could account for slightly larger deviations between results for the merged samples. However, we observe a reasonably low variation.

Figure 5 shows integrated rates of jets determined with the CDF Run II k_T-algorithm [65] as a function of the analysis cut d_{cut}. Monte Carlo results have been produced using a merged sample for up to five jets in the final state, generated with COMIX and showered with the CSS with a merging cut $Q_{\text{cut}} = 20\text{GeV}$.

Differential distributions

To study the merging systematics in more detail, we investigate again the differential jet rates $d_{n\,n+1}$, describing the scales at which $n+1$ jets are clustered into n jets according to the CDF Run II k_T-algorithm.

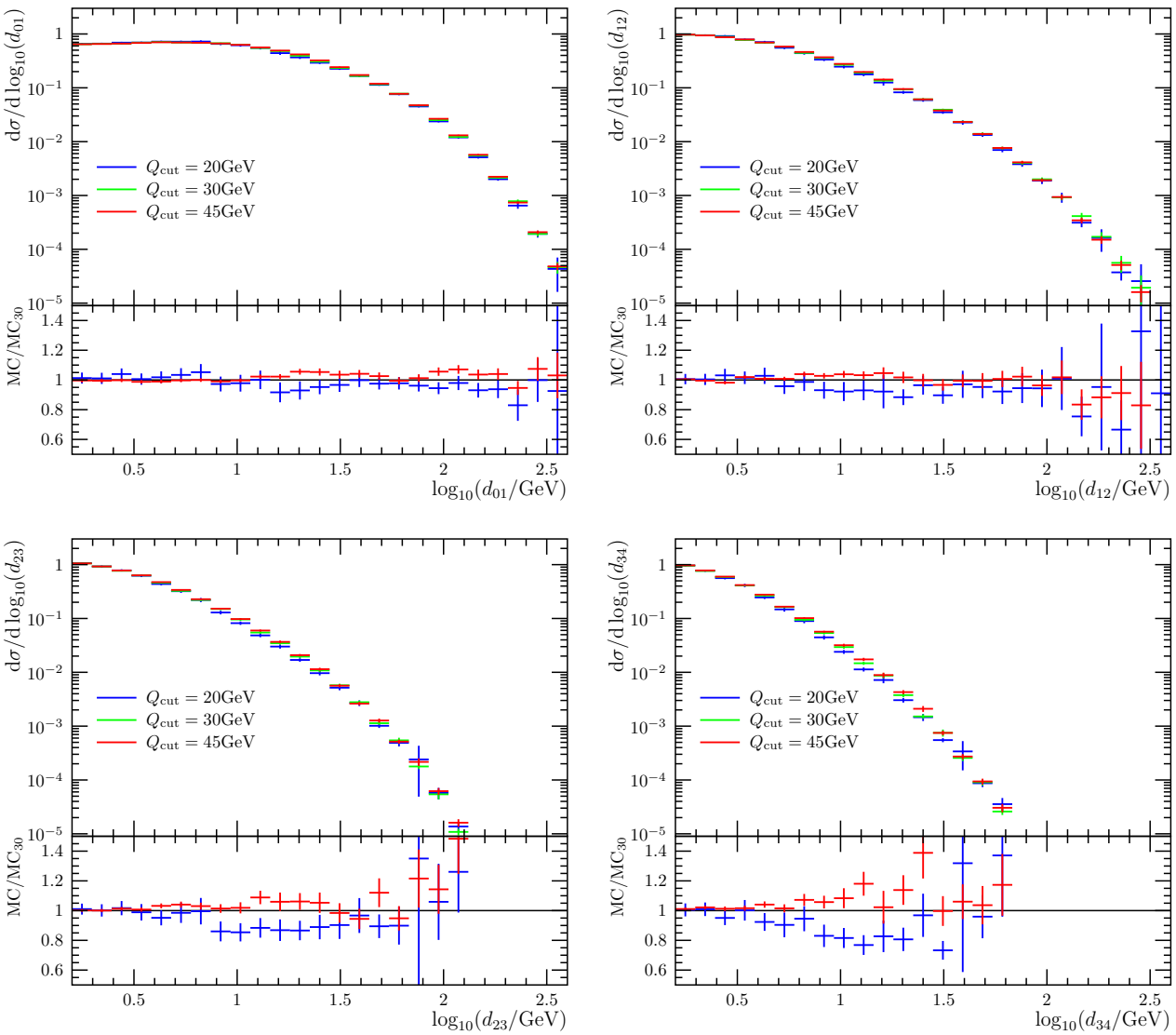

Figure 7: Differential jet rates $d_{n\,n+1}$ for three different merging cuts.

This algorithm has a free parameter, D, which accounts for the missing information on beam partons. Hence, in this setup, no firm relation can be established between the jet measure of the k_T-algorithm and the jet criterion, Eq. (24). Nevertheless, a certain correspondence between the two quantities exists, making these distributions a good testing ground for variations around the merging cuts.

To produce Fig. 6 a merged sample of up to five jets from the matrix element has been generated with COMIX and showered with the CSS. The merging cuts, which have been used, are $Q_{\rm cut} = 20\,\text{GeV}$, $Q_{\rm cut} = 30\,\text{GeV}$, and $Q_{\rm cut} = 45\,\text{GeV}$. As in the case of e^+e^- collisions, the deviations between the predictions of the various samples are small.

Most observables are even less sensitive to the precise value of the merging cut. As an example, Figure 7 displays the transverse momentum of the two leading jets for the three merging cut values in comparison to data from CDF [66].

It is also interesting to understand the influence of the maximal number of jets generated from the matrix element, $N_{\rm max}$, on experimental observables. We observe that typically the predictions are fairly stable for the $N_{\rm max}$ leading jets. To put it another way, for a given analysis investigating the n'th jet, one should use a Monte Carlo sample with $N_{\rm max} \geq n$. Due to the increased phase space available for QCD radiation at the LHC, the higher jet multiplicities will play an even more important role there.

Again, comparing to data from CDF [66] in Figure 8 and varying $N_{\rm max}$ between zero and three, the im-

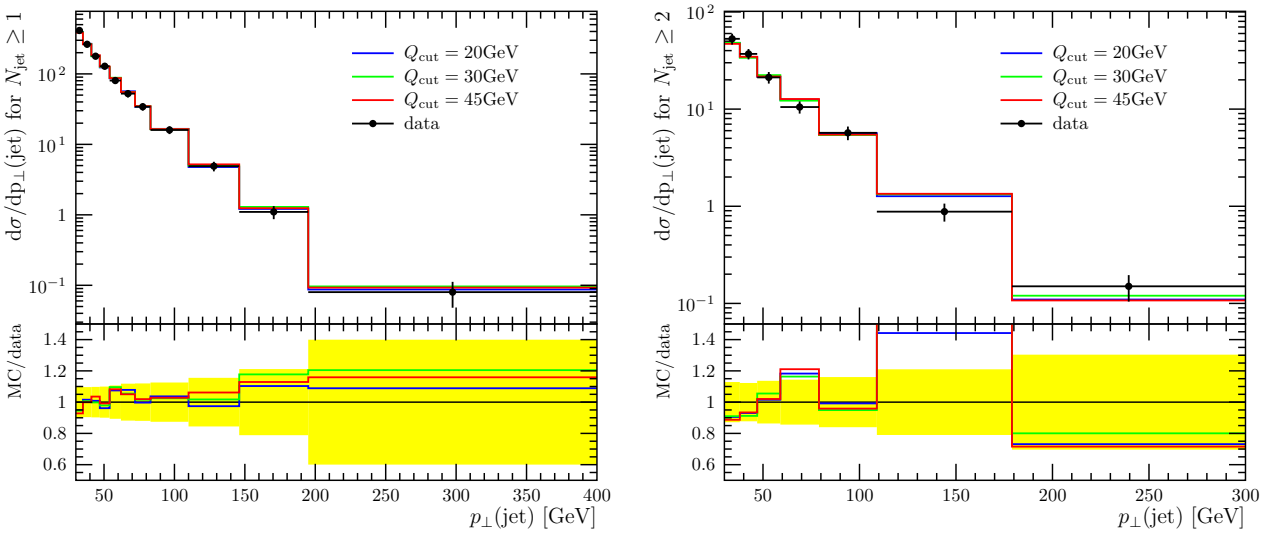

Figure 8: Jet p_T in $N_{\mathrm{jet}} \geq 1$ and $N_{\mathrm{jet}} \geq 2$ events compared to data from CDF [66].

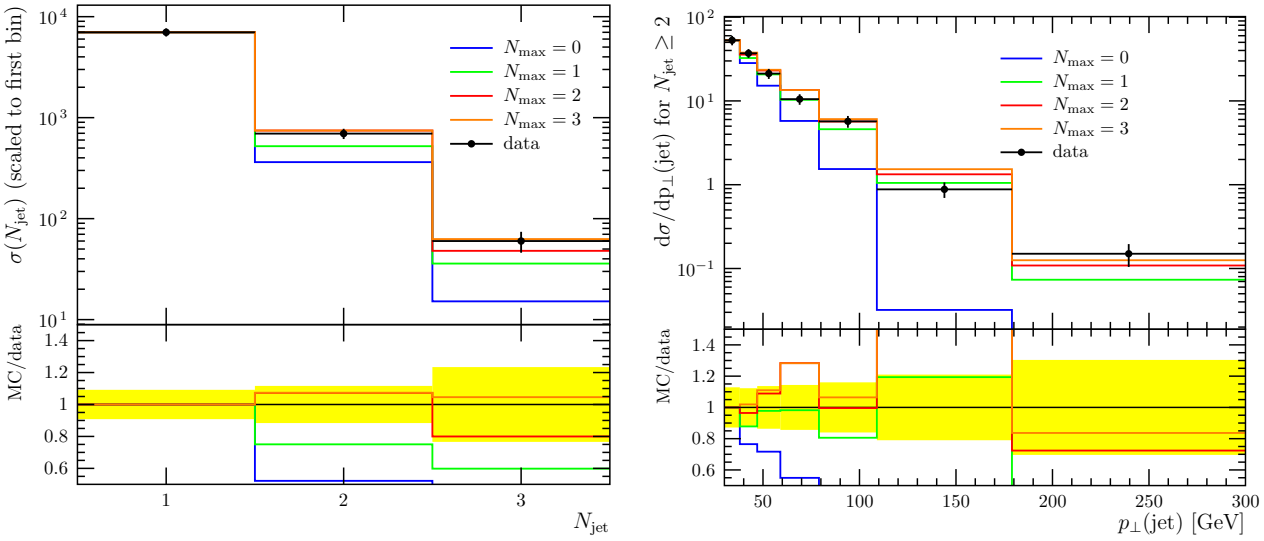

Figure 9: Jet multiplicity and jet p_T in $N_{\mathrm{jet}} \geq 2$ events compared to data from CDF [66].

portance of correctly describing additional hard jet production by the respective matrix elements can be estimated.

7 Conclusions

In this publication we have presented a general formal framework to discuss algorithms for the merging of multi-jet matrix elements and parton showers. We have constructed a merging algorithm that maintains the logarithmic accuracy provided by the parton shower in both initial and final state radiation. In this construction, special emphasis is put on an invariant formulation of the respective phase-space separation criterion. Because this criterion is not identical with the parton-shower evolution parameter, the logarithmic accuracy can only be maintained by running a truncated shower.

Hard matrix elements must be interpreted in the large-N_C limit to provide an input for shower Monte Carlos. Since the respective strategy is not unambiguous, the influence of different methods to assign colours was studied. We find no significant difference between the proposed algorithms, which range from heuristic assignment to the choice of a configuration with probability proportional to the respective colour ordered

subamplitude squared.

We have checked the systematics of the newly proposed algorithm in e^+e^- annihilation into hadrons and in Drell-Yan lepton pair production. We find greatly reduced uncertainties, compared to previously employed methods like the original CKKW technique. This statement holds for inclusive quantities such as total cross sections and jet rates, as well as for differential distributions.

Acknowledgements

We would like to thank Jan Winter, Tanju Gleisberg and Marek Schönherr for collaboration on the SHERPA project, which provided the framework for our study. Special thanks go to them as well as Thomas Binoth and Nigel Glover for fruitful discussions and their comments on the manuscript. We are grateful to Hendrik Hoeth and Andy Buckley for their help with RIVET, which we used for the analyses and the comparison with data. SH acknowledges funding by the HEPTOOLS Marie Curie Research Training Network (contract number MRTN-CT-2006-035505) and the Swiss National Science Foundation (SNF, contract number 200020-117602). FS gratefully acknowledges financial support by the MCnet Marie Curie Research Training Network (contract number MRTN-CT-2006-035606). The work of SS was partially funded by the STFC.

A Relation to other merging methods

In this appendix we discuss, on formal grounds, the relation of the merging algorithm proposed here with the original CKKW method [6]. Its relation with the CKKW-L method, presented in [13], and with the MLM method [15] will also briefly be discussed, arguing on more algorithmic grounds.

A.1 Formal equivalence with the CKKW method at NLL accuracy

Let us start by showing the formal equivalence, at next-to-leading logarithmic accuracy, of the CKKW method presented in [6] and the proposed merging algorithm. For this purpose, the generating function $\phi_a\left(Q,q\,;\phi^{(0)}\right)$ for a jet at scale q, produced by a jet that emerged at scale Q with initial condition $\phi^{(0)}$ is computed using the coherent-branching formalism. We focus on massless partons. All jet scales are given in terms of the Durham jet measure, which is defined as [10]

$$Q_{ij}^2 = 2\min\left\{E_i^2,E_j^2\right\}\,(1-\cos\theta_{ij}) = 2\,p_ip_j\,\min\left\{\frac{\tilde z_{i,j}}{1-\tilde z_{i,j}},\frac{1-\tilde z_{i,j}}{\tilde z_{i,j}}\right\}\,,\quad \text{where}\quad \tilde z_{i,j}=\frac{E_i}{E_i+E_j}\,. \tag{40}$$

Note that in the collinear region and for massless particles, the jet criterion, Eq. (24), effectively encodes the Durham measure for all splittings which are singular in $\tilde z_{i,j}$ or $1-\tilde z_{i,j}$. It is therefore justified to employ the Durham measure instead throughout this computation.

Collinear factorisation is assumed while the evolution parameter is the angular variable

$$t\to\tilde q^{\,2}=\frac{k_\perp^2}{z^2(1-z)^2}=\frac{p^2}{z(1-z)}\,, \tag{41}$$

with z the light-cone momentum fraction in the splitting and p^2 the virtuality of the mother parton.

Comparing the coherent-branching formalism to an arbitrary shower algorithm respecting colour coherence, potential differences arise because coherence is implemented through angular ordering. However, with regard to QCD evolution in the collinear regime, both methods are formally equivalent.

The generating function is defined by the differential equation

$$\begin{aligned}\frac{\mathrm{d}\phi_a\left(q,Q_0;\phi^{(0)}\right)}{\mathrm{d}\log(q/Q_0)} \;=\; &\int_0^1\mathrm{d}z\,\sum_{b=q,g}\mathcal{K}_{ab}(z,q)\,\Theta\left[\min\left\{z,1-z\right\}q-Q_0\right]\\ &\times\left[\phi_b\left(zq,Q_0;\phi^{(0)}\right)\,\phi_c\left((1-z)q,Q_0;\phi^{(0)}\right)-\phi_a\left(q,Q_0;\phi^{(0)}\right)\right]\,.\end{aligned} \tag{42}$$

where

$$\mathcal{K}_{ab}(z,q)\;=\;\frac{\alpha_s(z(1-z)\,q)}{2\pi}\,P_{a\to bc}(z)\,. \tag{43}$$

The upper bounds zq and $(1-z)q$ in ϕ_b and ϕ_c account for the modified starting conditions in the evolution of the respective jet [67]. The functions $\phi^{(0)}$ represent the initial condition for the evolution.

Equation (42) can be cast into a more suitable form by employing Sudakov form factors in terms of the Durham jet measure, computed at next-to-leading logarithmic accuracy, cf. [10]. For the purpose of this analysis we include the finite terms, such that

$$\begin{aligned}\bar{\Delta}_q(q,Q) &= \exp\left\{-\int_q^Q \frac{\mathrm{d}\bar{q}}{\bar{q}} \int_0^{1-q/Q} \mathrm{d}z\, 2\,\mathcal{K}_{qq}(z,q)\right\} \\ \bar{\Delta}_g(q,Q) &= \exp\left\{-\int_q^Q \frac{\mathrm{d}\bar{q}}{\bar{q}} \left[\int_{q/Q}^{1-q/Q} \mathrm{d}z\, \mathcal{K}_{gg}(z,q) + N_f \int_0^1 \mathrm{d}z\, \mathcal{K}_{gq}(z,q)\right]\right\} .\end{aligned} \tag{44}$$

Thus,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\log(q/Q_0)} \frac{\phi_a\left(q,Q_0;\phi^{(0)}\right)}{\bar{\Delta}_a(Q_0,q)} &= \frac{1}{\bar{\Delta}_a(Q_0,q)} \int_0^1 \mathrm{d}z \sum_{b=q,g} \mathcal{K}_{ab}(z,q) \\ &\times \Theta\left[\min\{z,1-z\}\, q - Q_0\right] \phi_b\left(zq,Q_0;\phi^{(0)}\right) \phi_c\left((1-z)q,Q_0;\phi^{(0)}\right) .\end{aligned} \tag{45}$$

For quark jets, employing that the largest contribution to the branching $q \to q$ comes from the region $1-z \ll 1$,

$$\frac{\mathrm{d}}{\mathrm{d}\log(q/Q_0)} \log\frac{\phi_q\left(q,Q_0;\phi^{(0)}\right)}{\bar{\Delta}_q(Q_0,q)} \approx \int_0^{1-Q_0/q} \mathrm{d}z\, 2\,\mathcal{K}_{qq}(z,q)\, \phi_g\left((1-z)q,Q_0;\phi^{(0)}\right) . \tag{46}$$

Note that the lower limit on the z integration is redefined to equal zero, because the largest contribution to the integral arises from the region $z \approx 1$. Using $\phi_q\left(Q_0,Q_0;\phi^{(0)}\right) = \phi_q^{(0)}$ and inserting the quark Sudakov form factor, yields

$$\phi_q\left(Q,Q_0;\phi^{(0)}\right) \approx \phi_q^{(0)} \exp\left\{\int_{Q_0}^Q \mathrm{d}q \int_0^{1-q/Q} \mathrm{d}z\, 2\,\mathcal{K}_{qq}(z,q) \left[\phi_g\left(q,Q_0;\phi^{(0)}\right) - 1\right]\right\} . \tag{47}$$

The solution to the respective evolution equation for gluon jets can be found in [10].

As discussed in [6], a vetoed parton shower can be expressed in terms of a different starting condition, $\phi^{(0)} \to \tilde{\phi}(Q,Q_{\mathrm{cut}},Q_0;\phi^{(0)})$. Then, in order to generate the correct jet fractions at some scale Q_0, the following identity is imposed:

$$\phi_a(Q,Q_{\mathrm{cut}};\tilde{\phi}) \stackrel{!}{=} \phi_a\left(Q,Q_0;\phi^{(0)}\right) . \tag{48}$$

Substituting the generating function ϕ_q, Eq. (47), yields, for the quark generating function,

$$\begin{aligned}\phi_q\left(Q,Q_0;\phi^{(0)}\right) &= \tilde{\phi}_q(Q,Q_{\mathrm{cut}},Q_0;\phi^{(0)}) \\ &\times \exp\left\{\int_{Q_{\mathrm{cut}}}^Q \mathrm{d}q \int_0^{1-q/Q} \mathrm{d}z\, 2\,\mathcal{K}_{qq}(z,q) \left[\phi_g\left(q,Q_0;\phi^{(0)}\right) - 1\right]\right\} .\end{aligned} \tag{49}$$

Consequently, the result for the modified parton shower reads

$$\tilde{\phi}_q(Q,Q_{\mathrm{cut}},Q_0;\phi^{(0)}) = \phi^{(0)} \exp\left\{\int_{Q_0}^{Q_{\mathrm{cut}}} \mathrm{d}q \int_0^{1-q/Q} \mathrm{d}z\, 2\,\mathcal{K}_{qq}(z,q) \left[\phi_g\left(q,Q_0;\phi^{(0)}\right) - 1\right]\right\} . \tag{50}$$

At next-to-leading logarithmic accuracy this is equivalent to the phase-space slicing method introduced in Sec. 3, because Eqs. (49) and (50) can be written as

$$\begin{aligned}\phi_q\left(Q,Q_0;\phi^{(0)}\right) &= \tilde{\phi}_q(Q,Q_{\mathrm{cut}},Q_0;\phi^{(0)}) \\ &\times \exp\left\{\int_{Q_0}^Q \mathrm{d}q \int_0^{1-q/Q} \mathrm{d}z\, 2\,\mathcal{K}_{qq}^{\mathrm{ME}}(z,q) \left[\phi_g\left(q,Q_0;\phi^{(0)}\right) - 1\right]\right\}\end{aligned} \tag{51}$$

and

$$\tilde{\phi}_q(Q, Q_{\text{cut}}, Q_0; \phi^{(0)}) \;=\; \phi^{(0)} \exp\left\{ \int_{Q_0}^{Q} \mathrm{d}q \int_0^{1-q/Q} \mathrm{d}z \, 2\, \mathcal{K}_{qq}^{\text{PS}}(z,q) \, \left[\phi_g\left(q, Q_0; \phi^{(0)}\right) - 1 \right] \right\} . \tag{52}$$

Here, the combination of an extended integration range for q and the cutoff due to the Θ-functions contained in $\mathcal{K}^{\text{ME}}$ and $\mathcal{K}^{\text{PS}}$ generates exactly the same q-integral as in Eqs. (49) and (50). A similar proof holds for the gluon-jet generating function.

Note that the extended range for z-integration in Eq. (46) would have been generated in a natural way by the jet criterion, Eq. (24). For $q \to qg$ splittings the Θ-function corresponding to the one in Eq. (42) would read $\Theta\left[(1-z)q - Q_0\right]$, thus setting the integration boundaries $0 \le z \le 1 - q/Q_0$. In any case the contribution from the range $0 \le z \le q/Q_0$ is subleading.

Algorithmic differences

Arguing on more algorithmic grounds, the following differences between the CKKW method and the new merging prescription presented in this publication can be established:

- Usually, Sudakov form factors within the CKKW approach are implemented in the form of (semi-)analytic functions, see for example [8]. The accuracy of the method is then limited by the level of correspondence between these functions and the respective parton-shower result.
- Truncated showers, which, according to Sec. 2.4, are mandatory to establish formal equivalence of the evolution in the pure parton shower and the matrix-element improved Monte Carlo algorithm, are implemented in an approximate way. They are replaced by modified starting conditions for external partons, rather than emissions from internal lines. The net effect is, that (up to typically less important momentum reshuffling) no redefinition of matrix element kinematics is necessary.

A.2 Correspondence to the CKKW-L method

In order to see the relation of the algorithm proposed here with the CKKW-L prescription presented in [13], let us concentrate first on the case of pure final-state radiation. A prime example for this is e^+e^--annihilations into hadrons. In this case, CKKW-L, based on ARIADNE [30], the jet criterion and the (dipole) shower-evolution parameter are identical. In fact, jets are defined by the dipole transverse momentum $p_\perp$ also used in the shower. For a splitting $\tilde{i}\tilde{k} \to ijk$,

$$p_\perp^2 = \frac{s_{ij} s_{jk}}{s_{\tilde{i}\tilde{k}}} \, . \tag{53}$$

Because both quantities coincide, there is no need for a truncated shower: Whenever an emission is harder than that of a node given by the hard matrix element, i.e. at a larger $p_\perp$, then it will produce a new jet. Since such emissions lead to vetoing the event, no intermediate emissions, which are not giving rise to a jet, must be taken into account. In this case the algorithm for e^+e^- annihilations reduces to

- Relevant multi-jet cross sections for the process under consideration are calculated with the phase-space restriction $Q > Q_{\text{cut}}$. Strong couplings are computed such that they give an overestimate, which can be reweighted.

1. Events are generated according to the above defined cross sections with kinematics determined by the respective matrix elements.
2. The most probable shower history of the final state is determined through backward clustering, cf. Sec. 2.3. The clustering is guided by information from the matrix element, which means that only those shower histories may be identified, which have a corresponding Feynman diagram.
3. The event is accepted or rejected according to a kinematics-dependent weight, which corresponds to evaluating strong couplings in the shower scheme and computing the no-branching probability for each dipole. This is done by starting the shower from the scale at which the dipole emerged. If the first

emission generated by the shower is harder than the scale given by the next matrix-element generated node related to that dipole (dipole contains at least one internal leg), or harder than $Q_{\rm cut}$ (dipole made of two external legs), then the event is rejected.

It is apparent that this is in perfect agreement with the method proposed in this paper, if the additional complication of the truncated shower can be neglected, which becomes obsolete if the shower-evolution parameter and the jet criterion coincide.

Comparing our method with the CKKW-L method in case of hadron collisions is not that simple. In principle, the extension of the algorithm for the dipole shower presented above to the case of hadron collisions is straightforward, especially with identical jet criterion and shower-evolution parameter. However, due to the fact that ARIADNE re-interprets initial-state radiation as final-state radiation with one or both dipole legs being the proton remnant, it is not entirely clear, how PDF effects are accounted for. An obvious way out would be to replace this non-perturbative algorithm for initial-state radiation with a formulation that rests entirely on the grounds of perturbation theory and the factorisation theorems, like the one presented in [32]. This opens an arena for further interesting studies.

A.3 Relation with the MLM method

It is difficult to establish any formal correspondence between the method proposed here and the MLM prescription [15]. This is mainly due to the fact that they indeed base on different ideas and methods. To our understanding, though, the main difference lies in the treatment of radiation off intermediate legs. Therefore we will very briefly describe the MLM method here and we will outline the apparent differences with respect to the new approach.

The MLM prescription, also presented in [18] bases on using a simple cone definition as the jet criterion to generate the parton configurations at the matrix-element level. The scales of strong couplings are reconstructed by using a backward clustering with a $k_\perp$ measure. Then the accepted configurations are passed on to parton-shower routines, typically the ones of PYTHIA [68] (virtuality or $p_\perp$-ordered shower) or of HERWIG [53] (angular-ordered shower). These codes typically reconstruct the parton-shower starting scales of a multi-parton configuration by directly inspecting colour connections, without any backward clustering and therefore partially neglecting intermediate legs and the radiation originating from them. Having the starting conditions at hand, the shower is invoked without any constraint. After it has terminated at the hadronisation scale the original partons stemming from the matrix element are matched to the jets present at parton-shower level, again defined by the cone algorithm. If such a match is not possible, either due to extra, unwanted jets being produced in the parton shower or due to "loosing" jets in the parton shower, the event is rejected.

The differences to the other algorithms are obvious:

- While the MLM prescription computes the Sudakov suppression weight in an inclusive way, the other algorithms, the original CKKW approach, the CKKW-L method and the algorithm proposed here, determine the Sudakov rejection by combining the rejection weights from the individual partons. Therefore, the effect of "loosing" jets is not present, and only the emergence of unwanted jets leads to vetoing the event.

- Because the definition of starting scales for the parton shower is left to the routines provided by PYTHIA and HERWIG it is not entirely clear, how far radiation of intermediate legs is accounted for.

  In any case, however, at present, neither of the two codes, PYTHIA and HERWIG allows truncated showering. This, following our argument, would be mandatory to formally guarantee the logarithmic accuracy provided by the respective parton-shower algorithm.

Despite these differences, a good agreement of predictions obtained with the CKKW method and the MLM method can often be observed. Respective results have been reported for example in [18]. They indicate, in particular, that the differences between the two methods may not be too important in the case of W+jets production at the Tevatron and the LHC. It can, however, not be expected that this statement holds for arbitrary processes.

B Merging with multiple leading-order processes

In this appendix, we consider situations like Drell-Yan production at large centre-of-mass energies. In such configurations the scale of the leading-order process – in our case the lepton pair production process – is low compared to the potential scales of additional jets. Hence, simply applying the pure merging prescription introduced in Sec. 3 will eventually result in missing hard radiation and therefore not lead to a good description of experimentally relevant signatures. An example for such signatures would be a lepton pair with mass lower than the transverse momentum of an accompanying QCD jet.

> *Note, that this is consistent with the proposed merging approach, since the choice of a specific factorisation scale at the leading order corresponds to generating additional QCD radiation up to an evolution parameter limited by this scale. Instead, the solution lies in the dynamic definition of a leading-order process for each event.*

Let us stick to the example of Drell-Yan production. Shower branchings cannot take place at scales larger than the factorisation scale, in this case usually the mass of the Drell-Yan pair, m_{ll}. Jets harder than this scale should thus actually be described by a leading-order process corresponding to "Drell-Yan + jet"-production. However, one needs to preserve the integrity of the inclusive Drell-Yan sample up to its factorisation scale.

The event-generation algorithm allows to do so. The central idea is to reinterpret the leading-order process as soon as the scale of jet production in any matrix-element configuration with additional jets exceeds the factorisation scale. The corresponding event is then not taken into account for the inclusive Drell-Yan production sample, but is internally treated as part of the "Drell-Yan + hard jet"-regime. Its factorisation scale and the corresponding shower starting conditions are redefined accordingly.

Note that the problem discussed in this appendix only occurs for leading-order configurations which do not include any strong coupling[7]. Also, jets at scales lower than the mass of the Drell-Yan pair are always well described in the merging. Such configurations can be produced by the parton shower, in principle at arbitrary multiplicity, because of the highest multiplicity treatment introduced in Sec. 3.3.

The relevance of the problem and the difference between the pure merging approach and the refined approach with multiple leading-order processes depends on the scales which are involved. If the phase space available to additional parton emission is not too large, like in the case of W or Z production at the Tevatron, resummation effects are small, because the difference between factorisation scale and the scale of hard extra jet production is usually small. In this case, the pure merging algorithm can safely be employed, since not too many hard jets can be generated, and this limited hard extra-jet multiplicity can be accounted for by matrix elements.

References

[1] M. H. Seymour, *Matrix-element corrections to parton shower algorithms*, Comp. Phys. Commun. **90** (1995), 95–101, [hep-ph/9410414]. G. Corcella and M. H. Seymour, *Matrix element corrections to parton shower simulations of heavy quark decay*, Phys. Lett. **B442** (1998), 417–426, [hep-ph/9809451]. E. Norrbin and T. Sjöstrand, *QCD radiation off heavy particles*, Nucl. Phys. **B603** (2001), 297–342, [hep-ph/0010012].

[2] S. Frixione and B. R. Webber, *Matching NLO QCD computations and parton shower simulations*, JHEP **06** (2002), 029, [hep-ph/0204244]. S. Frixione and B. R. Webber, *The MC@NLO 3.3 Event Generator*, hep-ph/0612272.

[3] S. Frixione, P. Nason and B. R. Webber, *Matching NLO QCD and parton showers in heavy flavour production*, JHEP **08** (2003), 007, [hep-ph/0305252]. S. Frixione, E. Laenen, P. Motylinski and B. R. Webber, *Single-top production in MC@NLO*, JHEP **03** (2006), 092, [hep-ph/0512250]. S. Frixione, E. Laenen, P. Motylinski, B. R. Webber and C. D. White, *Single-top hadroproduction in association with a W boson*, JHEP **07** (2008), 029, [arXiv:0805.3067 [hep-ph]].

[4] P. Nason, *A new method for combining NLO QCD with shower Monte Carlo algorithms*, JHEP **11** (2004), 040, [hep-ph/0409146]. S. Frixione, P. Nason and C. Oleari, *Matching NLO QCD computations*

[7] The case of effective theories, where QCD partons couple to non-QCD particles via loop graphs, which are in turn reinterpreted as effective vertices, is not considered here.

with parton shower simulations: the POWHEG method, JHEP **11** (2007), 070, [arXiv:0709.2092 [hep-ph]].

[5] P. Nason and G. Ridolfi, *A positive-weight next-to-leading-order Monte Carlo for Z pair hadroproduction*, JHEP **08** (2006), 077, [hep-ph/0606275]. O. Latunde-Dada, S. Gieseke and B. Webber, *A positive-weight next-to-leading-order Monte Carlo for e^+e^- annihilation to hadrons*, JHEP **02** (2007), 051, [hep-ph/0612281]. K. Hamilton, P. Richardson and J. Tully, *A Positive-Weight Next-to-Leading Order Monte Carlo Simulation of Drell-Yan Vector Boson Production*, arXiv:0806.0290 [hep-ph]. O. Latunde-Dada, *Applying the POWHEG method to top pair production and decays at the ILC*, arXiv:0806.4560 [hep-ph].

[6] S. Catani, F. Krauss, R. Kuhn and B. R. Webber, *QCD matrix elements + parton showers*, JHEP **11** (2001), 063, [hep-ph/0109231].

[7] F. Krauss, *Matrix elements and parton showers in hadronic interactions*, JHEP **0208** (2002), 015, [hep-ph/0205283].

[8] F. Krauss, A. Schälicke, S. Schumann and G. Soff, *Simulating W / Z + jets production at the Tevatron*, Phys. Rev. **D70** (2004), 114009, [hep-ph/0409106].

[9] F. Krauss, A. Schälicke, S. Schumann and G. Soff, *Simulating W/Z + jets production at the CERN LHC*, Phys. Rev. **D72** (2005), 054017, [hep-ph/0503280]. T. Gleisberg, F. Krauss, A. Schälicke, S. Schumann and J.-C. Winter, *Studying W^+W^- production at the Fermilab Tevatron with SHERPA*, Phys. Rev. **D72** (2005), 034028, [hep-ph/0504032].

[10] N. Brown and W. J. Stirling, *Finding Jets and Summing Soft Gluons: a New Algorithm*, Z. Phys. **C53** (1992), 629–636. S. Catani, Y. L. Dokshitzer, M. Olsson, G. Turnock and B. R. Webber, *New clustering algorithm for multijet cross sections in e^+e^- annihilation*, Phys. Lett. **B269** (1991), 432–438.

[11] S. Catani, Y. L. Dokshitzer and B. R. Webber, *The $k_\perp$ clustering algorithm for jets in deep inelastic scattering and hadron collisions*, Phys. Lett. **B285** (1992), 291–299. S. Catani, Y. L. Dokshitzer, M. H. Seymour and B. R. Webber, *Longitudinally-invariant $k_\perp$-clustering algorithms for hadron–hadron collisions*, Nucl. Phys. **B406** (1993), 187–224.

[12] T. Gleisberg, S. Höche, F. Krauss, A. Schälicke, S. Schumann and J. Winter, *SHERPA 1.α, a proof-of-concept version*, JHEP **02** (2004), 056, [hep-ph/0311263]. T. Gleisberg, S. Höche, F. Krauss, M. Schönherr, S. Schumann, F. Siegert and J. Winter, *Event generation with SHERPA 1.1*, JHEP **02** (2009), 007, [arXiv:0811.4622 [hep-ph]].

[13] L. Lönnblad, *Correcting the colour-dipole cascade model with fixed order matrix elements*, JHEP **05** (2002), 046, [hep-ph/0112284]. N. Lavesson and L. Lönnblad, *W+jets matrix elements and the dipole cascade*, JHEP **07** (2005), 054, [hep-ph/0503293].

[14] N. Lavesson and L. Lönnblad, *Extending CKKW-merging to one-loop matrix elements*, JHEP **12** (2008), 070, [arXiv:0811.2912 [hep-ph]].

[15] M. L. Mangano, M. Moretti and R. Pittau, *Multijet matrix elements and shower evolution in hadronic collisions: $Wb\bar{b}$ + n-jets as a case study*, Nucl. Phys. **B632** (2002), 343–362, [hep-ph/0108069].

[16] M. L. Mangano, M. Moretti, F. Piccinini and M. Treccani, *Matching matrix elements and shower evolution for top-pair production in hadronic collisions*, JHEP **01** (2007), 013, [hep-ph/0611129].

[17] J. Alwall, S. de Visscher and F. Maltoni, *QCD radiation in the production of heavy colored particles at the LHC*, JHEP **02** (2009), 017, [arXiv:0810.5350 [hep-ph]].

[18] S. Höche et al., *Matching Parton Showers and Matrix Elements*, hep-ph/0602031. J. Alwall et al., *Comparative study of various algorithms for the merging of parton showers and matrix elements in hadronic collisions*, Eur. Phys. J. **C53** (2008), 473–500, [arXiv:0706.2569 [hep-ph]].

[19] N. Lavesson and L. Lönnblad, *Merging parton showers and matrix elements – back to basics*, JHEP **04** (2008), 085, [arXiv:0712.2966 [hep-ph]].

[20] C. W. Bauer, F. J. Tackmann and J. Thaler, *GenEvA. (I). a new framework for event generation*, JHEP **12** (2008), 010, [arXiv:0801.4026 [hep-ph]]. C. W. Bauer, F. J. Tackmann and J. Thaler, *GenEvA (II): a phase space generator from a reweighted parton shower*, JHEP **12** (2008), 011, [arXiv:0801.4028 [hep-ph]].

[21] Z. Nagy and D. E. Soper, *Parton showers with quantum interference*, JHEP **09** (2007), 114, [arXiv:0706.0017 [hep-ph]]. Z. Nagy and D. E. Soper, *Parton showers with quantum interference: leading color, spin averaged*, JHEP **03** (2008), 030, [arXiv:0801.1917 [hep-ph]]. Z. Nagy and D. E. Soper, *Parton showers with quantum interference: leading color, with spin*, JHEP **07** (2008), 025, [arXiv:0805.0216 [hep-ph]].

[22] R. K. Ellis, W. J. Stirling and B. R. Webber, *QCD and collider physics*, ed. 1, vol. 8, Cambridge Monogr. Part. Phys. Nucl. Phys. Cosmol., 1996.

[23] J. C. Collins and D. E. Soper, *Back-to-back jets in QCD*, Nucl. Phys. **B193** (1981), 381. J. C. Collins and D. E. Soper, *Parton distribution and decay functions*, Nucl. Phys. **B194** (1982), 445. J. C. Collins and D. E. Soper, *The theorems of perturbative QCD*, Ann. Rev. Nucl. Part. Sci. **37** (1987), 383–409. J. C. Collins, D. E. Soper and G. Sterman, *Factorization of hard processes in QCD*, Adv. Ser. Direct. High Energy Phys. **5** (1988), 1–91, [hep-ph/0409313].

[24] K. Konishi, A. Ukawa and G. Veneziano, *A simple algorithm for QCD jets*, Phys. Lett. **B78** (1978), 243. K. Konishi, A. Ukawa and G. Veneziano, *Jet calculus: A simple algorithm for resolving QCD jets*, Nucl. Phys. **B157** (1979), 45–107. J. Kalinowski, K. Konishi and T. R. Taylor, *Jet calculus beyond leading logarithms*, Nucl. Phys. **B181** (1981), 221.

[25] S. Catani and L. Trentadue, *Resummation of the QCD perturbative series for hard processes*, Nucl. Phys. **B327** (1989), 323. A. H. Mueller, *On the multiplicity of hadrons in QCD jets*, Phys. Lett. **B104** (1981), 161–164. B. I. Ermolaev and V. S. Fadin, *Log-Log Asymptotic Form of Exclusive Cross-Sections in Quantum Chromodynamics*, JETP Lett. **33** (1981), 269–272. A. Bassetto, M. Ciafaloni, G. Marchesini and A. H. Mueller, *Jet multiplicity and soft gluon factorization*, Nucl. Phys. **B207** (1982), 189. Y. L. Dokshitzer, V. S. Fadin and V. A. Khoze, *Double Logs of Perturbative QCD for Parton Jets and Soft Hadron Spectra*, Z. Phys. **C15** (1982), 325.

[26] D. Amati, R. Petronzio and G. Veneziano, *Relating hard QCD processes through universality of mass singularities*, Nucl. Phys. **B140** (1978), 54. D. Amati, R. Petronzio and G. Veneziano, *Relating hard QCD processes through universality of mass singularities. 2*, Nucl. Phys. **B146** (1978), 29–49.

[27] V. N. Gribov and L. N. Lipatov, *Deep inelastic e-p scattering in perturbation theory*, Sov. J. Nucl. Phys. **15** (1972), 438–450. L. N. Lipatov, *The parton model and perturbation theory*, Sov. J. Nucl. Phys. **20** (1975), 94–102. Y. L. Dokshitzer, *Calculation of the structure functions for deep inelastic scattering and e^+e^- annihilation by perturbation theory in quantum chromodynamics*, Sov. Phys. JETP **46** (1977), 641–653. G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, Nucl. Phys. **B126** (1977), 298–318.

[28] V. V. Sudakov, *Vertex parts at very high-energies in quantum electrodynamics*, Sov. Phys. JETP **3** (1956), 65–71. G. P. Korchemsky, *Sudakov form-factor in QCD*, Phys. Lett. **B220** (1989), 629.

[29] G. Gustafson, *Dual description of a confined colour field*, Phys. Lett. **B175** (1986), 453. G. Gustafson and U. Pettersson, *Dipole formulation of QCD cascades*, Nucl. Phys. **B306** (1988), 746.

[30] L. Lönnblad, *Ariadne version 4: A program for simulation of QCD cascades implementing the colour dipole model*, Comput. Phys. Commun. **71** (1992), 15–31.

[31] W. T. Giele, D. A. Kosower and P. Z. Skands, *A Simple shower and matching algorithm*, Phys. Rev. **D78** (2008), 014026, [arXiv:0707.3652 [hep-ph]].

[32] J.-C. Winter and F. Krauss, *Initial-state showering based on colour dipoles connected to incoming parton lines*, JHEP **07** (2008), 040, [arXiv:0712.3913 [hep-ph]].

[33] Z. Nagy and D. E. Soper, *Matching parton showers to NLO computations*, JHEP **10** (2005), 024, [hep-ph/0503053]. Z. Nagy and D. E. Soper, *A new parton shower algorithm: Shower evolution, matching at leading and next-to-leading order level*, hep-ph/0601021.

[34] M. Dinsdale, M. Ternick and S. Weinzierl, *Parton showers from the dipole formalism*, Phys. Rev. **D76** (2007), 094003, [arXiv:0709.1026 [hep-ph]].

[35] S. Schumann and F. Krauss, *A parton shower algorithm based on Catani-Seymour dipole factorisation*, JHEP **03** (2008), 038, [arXiv:0709.1027 [hep-ph]].

[36] F. Maltoni, K. Paul, T. Stelzer and S. Willenbrock, *Color-flow decomposition of QCD amplitudes*, Phys. Rev. **D67** (2003), 014026, [hep-ph/0209271].

[37] C. Duhr, S. Höche and F. Maltoni, *Color-dressed recursive relations for multi-parton amplitudes*, JHEP **08** (2006), 062, [hep-ph/0607057].

[38] M. L. Mangano, S. J. Parke and Z. Xu, *Duality and multi-gluon scattering*, Nucl. Phys. **B298** (1988), 653.

[39] F. Caravaglios, M. L. Mangano, M. Moretti and R. Pittau, *A new approach to multi-jet calculations in hadron collisions*, Nucl. Phys. **B539** (1999), 215–232, [hep-ph/9807570].

[40] T. Sjöstrand, *Pythia 5.7 and Jetset 7.4 Physics and Manual*, hep-ph/9508391.

[41] S. Gieseke, *Uncertainties of Sudakov form factors*, JHEP **01** (2005), 058, [hep-ph/0412342].

[42] A. Banfi, G. P. Salam and G. Zanderighi, *Infrared-safe definition of jet flavour*, Eur. Phys. J. **C47** (2006), 113–124, [hep-ph/0601139].

[43] S. Catani, S. Dittmaier, M. H. Seymour and Z. Trocsanyi, *The dipole formalism for next-to-leading order QCD calculations with massive partons*, Nucl. Phys. **B627** (2002), 189–265, [hep-ph/0201036].

[44] R. Pittau, *A simple method for multi-leg loop calculations*, Comput. Phys. Commun. **104** (1997), 23–36, [hep-ph/9607309]. R. Pittau, *A simple method for multi-leg loop calculations 2: a general algorithm*, Comput. Phys. Commun. **111** (1998), 48–52, [hep-ph/9712418].

[45] S. Catani, S. Dittmaier and Z. Trocsanyi, *One-loop singular behaviour of QCD and SUSY QCD amplitudes with massive partons*, Phys. Lett. **B500** (2001), 149–160, [hep-ph/0011222].

[46] G. Belanger et al., *Automatic calculations in high energy physics and* GRACE *at one-loop*, Phys. Rept. **430** (2006), 117–209, [hep-ph/0308080]. J. Alwall et al., *MadGraph/MadEvent v4: The new web generation*, JHEP **09** (2007), 028, [arXiv:0706.2334 [hep-ph]].

[47] A. Kanaki and C. G. Papadopoulos, *HELAC: A package to compute electroweak helicity amplitudes*, Comput. Phys. Commun. **132** (2000), 306–315, [hep-ph/0002082].

[48] M. L. Mangano, M. Moretti, F. Piccinini, R. Pittau and A. D. Polosa, *ALPGEN, a generator for hard multiparton processes in hadronic collisions*, JHEP **07** (2003), 001, [hep-ph/0206293]. W. Kilian, T. Ohl and J. Reuter, *WHIZARD: Simulating Multi-Particle Processes at LHC and ILC*, arXiv:0708.4233 [hep-ph].

[49] F. Krauss, R. Kuhn and G. Soff, *AMEGIC++ 1.0: A Matrix Element Generator In C++*, JHEP **02** (2002), 044, [hep-ph/0109036].

[50] T. Gleisberg and S. Höche, *Comix, a new matrix element generator*, JHEP **12** (2008), 039, [arXiv:0808.3674 [hep-ph]].

[51] T. Sjöstrand, L. Lönnblad, S. Mrenna and P. Skands, *PYTHIA 6.3 Physics and Manual*, hep-ph/0308153. R. Kuhn, F. Krauss, B. Ivanyi and G. Soff, *APACIC++ 1.0: A PArton Cascade In C++*, Comput. Phys. Commun. **134** (2001), 223–266, [hep-ph/0004270]. F. Krauss, A. Schälicke and G. Soff, *APACIC++ 2.0: A PArton Cascade In C++*, Comput. Phys. Commun. **174** (2006), 876–902, [hep-ph/0503087].

[52] T. Sjöstrand and P. Z. Skands, *Transverse-momentum-ordered showers and interleaved multiple interactions*, Eur. Phys. J. **C39** (2005), 129–154, [hep-ph/0408302].

[53] G. Corcella et al., *HERWIG 6.5 Release Note*, hep-ph/0210213. M. Bähr et al., *Herwig++ Physics and Manual*, Eur. Phys. J. **C58** (2008), 639–707, [arXiv:0803.0883 [hep-ph]].

[54] R. Kleiss and W. J. Stirling, *Spinor techniques for calculating $p\bar{p} \to W^{\pm}/Z^0$+jets*, Nucl. Phys. **B262** (1985), 235–262. A. Ballestrero and E. Maina, *A new method for helicity calculations*, Phys. Lett. **B350** (1995), 225–233, [hep-ph/9403244].

[55] S. Catani and M. H. Seymour, *A general algorithm for calculating jet cross sections in NLO QCD*, Nucl. Phys. **B485** (1997), 291–419, [hep-ph/9605323].

[56] Y. L. Dokshitzer and G. Marchesini, *Monte Carlo and large angle gluon radiation*, JHEP **03** (2009), 117, [arXiv:0809.1749 [hep-ph]].

[57] J. Pumplin, D. R. Stump, J. Huston, H. L. Lai, P. Nadolsky and W. K. Tung, *New generation of parton distributions with uncertainties from global QCD analysis*, JHEP **0207** (2002), 012, [hep-ph/0201195].

[58] J.-C. Winter, F. Krauss and G. Soff, *A modified cluster-hadronisation model*, Eur. Phys. J. **C36** (2004), 381–395, [hep-ph/0311085].

[59] F. Krauss, T. Laubrich and F. Siegert, *Simulation of hadron decays in SHERPA*, in preparation.

[60] M. Schönherr and F. Krauss, *Soft photon radiation in particle decays in SHERPA* , JHEP **12** (2008), 018, [arXiv:0810.5071 [hep-ph]].

[61] A. Buckley, H. Hoeth, H. Schulz and J. E. von Seggern, *Monte Carlo event generator validation and tuning for the LHC*, arXiv:0902.4403 [hep-ph].

[62] B. M. Waugh et al., *HZTool and Rivet: Toolkit and framework for the comparison of simulated final states and data at colliders*, hep-ph/0605034. A. Buckley, *Tools for event generator tuning and validation*, arXiv:0809.4638 [hep-ph].

[63] P. Pfeifenschneider et al., JADE and OPAL collaboration, *QCD analyses and determinations of α_s in e^+ e^- annihilation at energies between 35-GeV and 189-GeV*, Eur. Phys. J. **C17** (2000), 19–51, [hep-ex/0001055].

[64] P. Abreu et al., DELPHI collaboration, *Tuning and test of fragmentation models based on identified particles and precision event shape data*, Z. Phys. **C73** (1996), 11–60.

[65] G. C. Blazey et al., *Run II jet physics*, hep-ex/0005012.

[66] T. Aaltonen et al., CDF - Run II collaboration, *Measurement of Inclusive Jet Cross Sections in $Z/\gamma^*(\to ee)$+ jets Production in $p\bar{p}$ Collisions at $\sqrt{s} = 1.96$ TeV*, Phys. Rev. Lett. **100** (2008), 102001, [arXiv:0711.3717 [hep-ex]].

[67] S. Catani, B. R. Webber and G. Marchesini, *QCD coherent branching and semiinclusive processes at large x*, Nucl. Phys. **B349** (1991), 635–654.

[68] T. Sjöstrand, S. Mrenna and P. Skands, *PYTHIA 6.4 physics and manual*, JHEP **05** (2006), 026, [hep-ph/0603175].